浙江省重点专业资助项目

杭州市重点学科资助项目

JINSHI DAISHU GUANDIAN XIA DE GAODENG DAISHU

近世代数观点下的高等代数

陈辉 著

$$f(t)=\sum_{k=0}^{\infty}\frac{<f(t)\mid x^k>}{k!}t^k$$

$$f(t)g(t)=\sum_{m=0}^{\infty}\left(\frac{1}{m!}\sum_{k=0}^{m}\binom{m}{k}a_k b_{m-k}\right)t^m$$



ZHEJIANG UNIVERSITY PRESS

浙江大学出版社

- 本书在近世代数思想指导下对高等代数的基本概念、基础理论、基本方法进行系统归纳与提升，同时把国内外有关高等代数研究的新成果引入本书.

- 首先概括地介绍了高等代数的一些主要内容，包括多项式理论、矩阵理论、向量空间和线性变换、欧氏空间和二次型等基础理论.

- 详细讨论了近世代数的一些主要内容，包括群、环、域、模等代数系统，又进一步讨论了主理想整环上的模理论，证明了有限生成模的循环分解定理. 这一定理对于后面讨论的有限维线性算子的结构定理是至关重要的.

- 最后对代数学的后续内容进行了讨论，把这些内容归纳为几个专题：线性算子的结构理论、谱理论、赋范线性空间、希尔伯特空间、双线性映射与张量积、仿射几何与多项式函数等.

JINSHI DAISHU GUANDIAN XIA DE GAODENG DAISHU

ISBN 978-7-308-06882-6



9 787308 068826 >
定价：45.00元

■ 浙江省重点专业资助项目
杭州市重点学科资助项目

近世代数观点下的高等代数

陈 辉 著


ZHEJIANG UNIVERSITY PRESS
浙江大学出版社

图书在版编目(CIP)数据

近世代数观点下的高等代数/陈辉著. —杭州：浙江大学出版社，2009.8
浙江省重点专业资助项目、杭州市重点学科资助项目
ISBN 978-7-308-06882-6

I. 近… II. 陈… III. 高等代数—研究 IV. 015

中国版本图书馆 CIP 数据核字(2009)第 112940 号

近世代数观点下的高等代数

陈 辉 著

责任编辑 阮海潮(ruanhc@163.com)
封面设计 姚燕鸣
出版发行 浙江大学出版社
(杭州天目山路 148 号 邮政编码 310028)
(网址：http://www.zjupress.com)
排 版 杭州大漠照排印刷有限公司
印 刷 杭州杭新印务有限公司
开 本 787mm×960mm 1/16
印 张 20.5
字 数 436 千
版 印 次 2009 年 8 月第 1 版 2009 年 8 月第 1 次印刷
书 号 ISBN 978-7-308-06882-6
定 价 45.00 元

浙江大学出版社发行部邮购电话(0571)88925591

前　　言

代数学是近代数学的一个重要分支，随着科学的发展和实际应用的需要，代数学的内容和方法都在不断地扩充和深化，特别是由于电子计算机的飞速发展和广泛应用，使许多实际问题可以通过离散化的数值计算得到定量的解决，这是离不开代数理论的.

为了适应新的形势，满足广大读者深入学习和研究的需求，作者根据多年的积累完成此书.我们是在近世代数思想指导下对线性代数的基本概念、基础理论、基本方法进行系统归纳与提升，同时把国内外线性代数研究的新成果引入本书.

本书首先概括地介绍了高等代数的一些主要内容，包括多项式理论、矩阵理论、向量空间和线性变换、欧氏空间和二次型等基础理论.

接下来讨论了近世代数的一些主要内容，包括群、环、域、模等代数系统，又进一步讨论了主理想整环上的模理论，证明了有限生成模的循环分解定理.这一定理对于后面讨论的有限维线性算子的结构定理是至关重要的.

最后对代数学的后续内容进行了讨论，把这些内容归纳为几个专题：线性算子的结构理论、谱理论、赋范线性空间、希尔伯特空间、双线性映射与张量积、仿射几何与多项式函数等.各章都是由各自相对独立的主题所组成的.

编写本书的目的是使读者能用近世代数观点来讨论高等代数问题，通过对这些专题的深入讨论，激发读者的原始性创新，从而培养读者的发散型思维.

陈　辉

2009 年 7 月

目　　录

第 1 章　基础知识

本章讲述的是全书经常会用到的一些基本概念,将主要介绍:集合与映射、代数系、等价关系、基数(势)的基本性质以及同态与同构等基本概念.通过本章的学习,将对近世代数常用的方法有初步的了解,为后续学习奠定基础.

1.1　集合与映射

1.1.1　集合

在数学中,经常讨论的不是孤立的个体,也不是包罗万象的宇宙,而往往是对具有某些特性之个体的联合体进行研讨,这样就产生了集合的概念.这是数学上一个不定义的原始概念,通常是用各种同义语来解释的.

定义 1.1.1　若干个确定事物的全体称为一个**集合**,简称为**集**.

把不含有任何元素的集合称为**空集**,记为 $\varnothing$.一个集合若仅含有有限个元素,则称该集合为**有限集**;若集合含有无限多个元素,则称该集合为**无限集**.

定义 1.1.2　集合的相等:如果集合 A 与集合 B 含有完全相同的元素,那么就称**集 A 与集 B 相等**,记为 $A = B$.

换言之:$A = B \Leftrightarrow A \subseteq B$ 且 $B \subseteq A \Leftrightarrow \forall x \in A$ 有 $x \in B$, $\forall y \in B$ 有 $y \in A$.这是判别集合相等的方法.

表示集合的方法一般有列举法和描述法.所谓列举法,就是把集合的所有元素直接列出来;所谓描述法,就是规定集合中元素所具有的性质.

两种方法都用符号$\{\cdots\}$来表示.列举法用$\{a,b,c\ \cdots\}$来表示;描述法用$\{x \mid p(x)\}$来表示,其中 $p(x)$ 表示元素 x 所具有的性质.

例如:$A = \{1, -1\}$,$B = \{1, 2, 3 \cdots 10\}$;$C = \{x \mid x \in \mathbf{R}, x^3 - 2 = 0\}$.

前者具体写出 A 是由 1 与 -1 两个整数组成的集合,B 是由 1 到 10 十个整数组成的集合;C 是方程 $x^3 - 2 = 0$ 的实数根所组成的集合.

常用数集:自然数集 $\mathbf{N}$,整数集 $\mathbf{Z}$,有理数集 $\mathbf{Q}$,实数集 $\mathbf{R}$,复数集 $\mathbf{C}$.

例 1.1.1　设 F 是数域,F 上所有一元多项式以及零多项式构成的集合记为:

$F[x] = \{a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \mid a_i \in F, i = 1, 2 \cdots n, n \in \mathbf{N}\}$.

例 1.1.2 数域 F 上所有 n 阶方阵的集合 $M_n(F) = \{(a_{ij})_n \mid a_{ij} \in F, i, j = 1, 2 \cdots n\}$.

定义 1.1.3 如果 B 的每个元素都属于 A，那么称 B 为 A 的**子集**，记为 $B \subseteq A$.

如果集合 B 是 A 的子集，且 $B \neq A$，则称 B 为 A 的**真子集**，记为 $B \subset A$.

$\varnothing$ 被认为是任意集合的子集，因此，对于任意集合 A，总有 $\varnothing \subseteq A$，$A \subseteq A$，并称其为 A 的**平凡子集**；A 的其他子集称为 A 的**非平凡子集**.

定义 1.1.4 设 A 是一个给定的集合，由 A 的所有子集组成的集合称为 A 的**幂集**(power set)，用符号 2^A 表示.

例 1.1.3 设 $A = \{0,1,2\}$，则 $2^A = \{\varnothing, \{0\}, \{1\}, \{2\}, \{0, 1\}, \{0, 2\}, \{1, 2\}, A\}$.

1.1.2 集合的运算

设 I 是一个集合，A, B, C 都是 I 的子集.

1) 由一切属于 A 或者属于 B 的元素组成的集合称为 A 与 B 的**并集**，记为 $A \cup B$，即 $A \cup B = \{x \mid x \in A$ 或 $x \in B\}$；

2) 由 A 与 B 公共的元素组成的集合称为 A 与 B 的**交集**，记为 $A \cap B$，即 $A \cap B = \{x \mid x \in A$ 且 $x \in B\}$；

3) 由一切属于 A 但不属于 B 的元素组成的集合称为 A 与 B 的**差集**，记为 $A - B$，即 $A - B = \{x \mid x \in A$ 且 $x \in B\}$；

4) 由一切不属于 A(但属于 I) 的元素组成的集合，称为 A 的**余集**，记为 A'，即 $A' = I - A$.

这些运算满足以下运算规律：

(1) $A \cup A = A \cap A = A$ （幂等律）

(2) $A \cup B = B \cup A, A \cap B = B \cap A$ （交换律）

(3) $A \cup (B \cup C) = (A \cup B) \cup C$；$A \cap (B \cap C) = (A \cap B) \cap C$ （结合律）

(4) $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$，（分配律）

$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$

(5) $A \cap (A \cup B) = A \cup (A \cap B) = A$ （吸收律）

(6) 若 $A \subseteq C$，则 $A \cup (B \cap C) = (A \cup B) \cap C$ （模　律）

(7) $(A \cup B)' = A' \cap B'$，$(A \cap B)' = A' \cup B'$ （DeMorgan 律）

(8) $(A')' = A$ （对合律）

这些运算与运算规律可推广到多个子集的情形.

设 $A_1, A_2 \cdots A_n$ 是 n 个任意集合，它们的并集记为 $\cup A_i$，定义为：

$$\bigcup A_i = A_1 \cup A_2 \cup \cdots \cup A_n = \{x \mid x \in A_i, \text{对于某一个 } i\}$$

这 n 个集合的交集记为 $\bigcap A_i$,定义为:

$$\bigcap A_i = A_1 \cap A_2 \cap \cdots \cap A_n = \{x \mid x \in A_i, i = 1, 2 \cdots n\}.$$

定义 1.1.5　设 A,B 是两个集合,由一切有序对 (a, b),其中 $a \in A$,$b \in B$ 组成的集合,称为 A 与 B 的**乘积集**,记为 $A \times B$,即 $A \times B = \{(a, b) \mid a \in A, b \in B\}$,并且规定,$(a_1, b_1) = (a_2, b_2) \Leftrightarrow a_1 = a_2, b_1 = b_2$.

例 1.1.4　笛卡尔坐标平面上点的坐标的集合就是 $\mathbf{R} \times \mathbf{R} = \{(a, b) \mid a, b \in \mathbf{R}\}$.

又,$\mathbf{R} \times \{1\} = \{(y, 1) \mid y \in \mathbf{R}\}$,显然是由过点 $(0, 1)$ 且平行于 x 轴的直线上的所有点的坐标组成的集合.

乘积集的概念也可以推广到多个集合的情形,设 $A_1, A_2 \cdots A_n$ 是任意 n 个集合,令

$$A_1 \times A_2 \times \cdots \times A_n = \{(a_1, a_2 \cdots a_n) \mid a_i \in A_i, i = 1, 2 \cdots n\}$$

其中

$$(a_1, a_2 \cdots a_n) = (b_1, b_2 \cdots b_n) \Leftrightarrow a_i = b_i,\ i = 1, 2 \cdots n,$$

称 $A_1 \times A_2 \times \cdots \times A_n$ 为**乘积集**,也可以记为 $\prod_{i=1}^{n} A_i$. 如果 $A_1, A_2 \cdots A_n$ 均为有限,那么 $\left|\prod_{i=1}^{n} A_i\right| = \prod_{i=1}^{n} |A_i|$.

1.1.3　包含与排斥原理

关于集合运算后元素个数的变化有以下规律:设 I 是一个集合,A,B,C 是 I 的有限子集,则有

$|A \cup B| = |A| + |B| - |A \cap B|$,

$|A \cap B| = |A| + |B| - |A \cup B|$,

$|A \cap B \cap C| = |A| + |B| + |C| - |A \cup B| - |A \cup C| - |B \cup C| + |A \cup B \cup C|$,

$|A \cup B \cup C| = |A| + |B| + |C| - |A \cap B| - |A \cap C| - |B \cap C| + |A \cap B \cap C|$.

当 $A \cap B = \varnothing$ 时,有 $|A \cup B| = |A| + |B|$,这就是"加法原理". 这些公式很容易用图形加以证明. 对于多个子集的情形有以下定理:

定理 1.1.1(包含与排斥原理)　设 $A_1, A_2 \cdots A_n$ 是 I 的有限子集,则

$$\left|\bigcap_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} |A_i| - \sum_{1 \leqslant i < j \leqslant n} |A_i \cup A_j| + \cdots + (-1)^{n-1} \left|\bigcup_{i=1}^{n} A_i\right|, \tag{1.1}$$

$$\left|\bigcup_{i=1}^{n} A_i\right| = \sum_{i=1}^{n} |A_i| - \sum_{1 \leqslant i < j \leqslant n} |A_i \cup A_j| + \cdots + (-1)^{n-1} \left|\bigcap_{i=1}^{n} A_i\right|. \tag{1.2}$$

下面举例说明包含与排斥原理的应用.

例 1.1.5 求不大于 500 可被 5,7,9 中某一个数整除的正整数的个数.

解 设不大于 500 可被 5 整除的正整数集合为 A_1,不大于 500 可被 7 整除的正整数集合为 A_2,不大于 500 可被 9 整除的正整数集合为 A_3,则

$$|A_1| = 100, |A_2| = [500/7] = 71, |A_3| = [500/9] = 55.$$

$$|A_1 \cap A_2| = [500/35] = 14, \quad |A_1 \cap A_3| = [500/45] = 11,$$

$$|A_2 \cap A_3| = [500/63] = 7, \quad |A_1 \cap A_2 \cap A_3| = [500/315] = 1.$$

故由(2)式得:

$$\begin{aligned} |A_1 \cup A_2 \cup A_3| &= \sum |A_i| - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| + |A_1 \cap A_2 \cap A_3| \\ &= 100 + 71 + 55 - 14 - 11 - 7 + 1 \\ &= 195. \end{aligned}$$

1.1.4 映射

代数研究的基本对象是代数系统,所谓代数系统就是带有运算的集合.这些代数系统的内部以及互相之间往往存在一定的联系,而映射则是建立这种联系的有用工具,所以说映射概念是近代数学最基本的概念.

映射是函数概念的推广,它描述了两个集合的元素之间的关系,是数学中最基本的工具之一,我们必须对它十分熟悉.

定义 1.1.6 设 A,B 是两个给定的集合,若存在一个 A 到 B 的对应关系 f,使得对 A 中的每一个元素 x,都有 B 中唯一确定的元素 y 与之对应,则称 f 是 A 到 B 的一个**映射**,记作:$y = f(x)$.其中 y 称为 x 的**象**(image),x 称为 y 的**原象**(inverse image),A 称为 f 的**定义域**(domain),B 称为 f 的**值域**(codomain).

通常,用 $f: A \to B$ 抽象地表示 f 是 A 到 B 的一个映射,用记号 $f: x \to y$ 表示映射 f 所规定的元素之间的具体对应关系.

例 1.1.6 设 $A = \{a, b, c, d\}$,$B = \{1, 2, 3, 4\}$. $f: a \to 1, b \to 2, c \to 4, d \to 4$,则 f 满足定义中的条件,是一个 A 到 B 的映射.

例 1.1.7 设 $A = \{1,2\}$,$B = \mathbf{Z}$,规定 A 到 B 的对应为 $f: 1 \to$ 奇数,$2 \to$ 偶数.

由于 $\mathbf{Z}$ 中的奇、偶数不止一个,故 $f(1)$,$f(2)$ 都不唯一确定,故 f 不是 A 到 B 的映射.

主要是由于自变量的表达形式不唯一而引起象的不唯一.因此,遇到这种情况时检验一个对应关系 f 是否是映射需检验是否满足下列条件:若 $x_1 = x_2$,则 $f(x_1) = f(x_2)$.

定义 1.1.7(映射的相等) 由于一个映射由定义域、值域、对应关系三个因素决定,所以两个映射相等必须这三个因素都相等,即如果 $f: A_1 \to B_1, g: A_2 \to B_2$,当且仅当 $A_1 = A_2, B_1 = B_2$,和任意 $x \in A_1$ 有 $f(x) = g(x)$ 时,称 f 与 g **相等**,记作 $f = g$.

例 1.1.8 设 $A = (-\infty, +\infty)$, $B = [0, +\infty)$,对于 $\forall x \in A$ 令

$$f: x \to |x| + 1,\ g: x \to \sqrt{x^2} + 1,\text{则 } f = g.$$

定义 1.1.8 设 f 是 A 到 B 的一个映射,称 $f(A) = \{f(x) \mid \forall x \in A\}$ 为 A 在 f 下的**象集**(也记为 $\mathrm{Im} f$),显然 $f(A)$ 是 B 的子集;

称 $f^{-1}(B) = \{x \in A \mid f(x) \in B\}$ 为 B 在 f 下的**完全原象集**,显然 $f^{-1}(B)$ 是 A 的子集.

定义 1.1.9 设 f 是 A 到 B 的一个映射,如果 $f(A) = B$,则称 f 是 A 到 B 的**满射**(surjection).

设 f 是 A 到 B 的一个映射,$\forall x_1$、$x_2 \in A$ 且 $x_1 \neq x_2$,若 $f(x_1) \neq f(x_2)$,则称 f 是 A 到 B 的一个**单射**(injection).

如果 f 既是满射又是单射,则称 f 是 A 到 B 的**一一映射**(bijection),或**双射**.

要证明 f 是满射,就是要证明 $f(A) = B$,也就是要证明 $f(A) \supseteq B$,即 $\forall b \in B$,存在 $a \in A$ 使得 $f(a) = b$. 这是判断 f 是不是满射的有效方法.

要证明 f 是单射,就是要证明它的逆否命题:如果 $f(x_1) = f(x_2)$,就有 $x_1 = x_2$.

定义 1.1.10 令 f 是 A 到 B 的一个映射,如果 $X \subset A$,那么 $f: A \to B$ 的**限制**是函数 $f|_X: X \to B$. 显然,单射的限制是单射.

例 1.1.9 设 $A = \{0,1,2,3\cdots\}$, $B = \{1,2,3\cdots\}$,定义对应关系:$f: n \to n+1$,不难验证 f 是一一映射.

例 1.1.10 设 $M = \{1/2, 1/3, 1/4\cdots\}$, $N = \{0, 1, 1/2, 1/3\cdots\}$,则 M 是 $(0, 1)$ 的子集,N 是 $[0, 1]$ 的子集. 作 $(0,1)$ 到 $[0,1]$ 的对应关系

$$f: 1/2 \to 0,\quad 1/3 \to 1,\quad 1/4 \to 1/2,\quad 1/5 \to 1/3\quad \cdots\quad 1/n \to 1/(n-2)\cdots$$

显然 f 是 $(0,1)$ 到 $[0,1]$ 的一一映射.

若 f 是 A 到 A 自身的映射,则称 f 是 A 上的一个**变换**.

例 1.1.11 $\forall f(x) \in F[x]$,规定 $\varphi: f(x) \to f'(x)$(这里表示导数),则 φ 是 $F[x]$ 的一个变换,并且是一个满变换.

当 A 是有限集时,A 上的变换通常用“列表法”表示.

1.1.5 映射的合成

类似于熟知的复合函数的概念,我们给出两个映射复合的概念.

定义 1.1.11 设 A,B,C 为三个集合,有两个映射:

若 $f: A \to B, g: B \to C$,则 f, g 可确定 A 到 C 的映射 $h: h(x) = g(f(x)), \forall x \in A$,称 h 是 f 与 g 的**合成**(或**复合**)(composite),记作:$h = g \circ f$.

设 I_A 是 A 上的一个变换,若对于 $\forall x \in A$,有 $I_A(x) = x$,则称 I_A 是 A 上的一个**单位变换**或**恒等变换**.映射的复合有时也称为映射的**乘法**,记为 $g \cdot f$ 或 gf.

根据复合映射的定义,f 的值域与 g 的定义域必须相同,才能复合成为 $g \circ f$.

对于 A 的任意两个变换 f, g 总是可以合成为 A 的变换 gf 或 fg,但一般 $gf \neq fg$.

关于映射的复合有以下性质:

定理 1.1.2 设有映射 $f: A \to B, g: B \to C, h: C \to D$,则有:

1) $h(gf) = (hg)f$;

2) $I_B f = f I_A = f$.

证明 1) 易见 $h(gf)$ 和 $(hg)f$ 的定义域都是 A,值域都是 D,且对于 $\forall x \in A$,有 $(h(gf))(x) = h((gf)(x)) = h(g(f(x))), ((hg)f)(x) = (hg)(f(x)) = h(g(f(x)))$,所以 $h(gf) = (hg)f$.

2) 若 $I_B f$ 和 f 的定义域都是 A,值域都是 B,且 $\forall x \in A$,则有 $I_B f(x) = I_B(f(x)) = f(x)$.

类似于反函数,对映射有逆映射的概念.

定义 1.1.12 设 $f: A \to B$,

1) 若存在映射 $g: B \to A$,使 $gf = I_A$,就称 g 是 f 的**左逆**;

2) 若存在映射 $h: B \to A$,使 $fh = I_B$,就称 h 是 f 的**右逆**;

3) 若 f 同时有左、右逆,则左、右逆相等,称为 f 的**逆**(inverse),记作 f^{-1}.

定理 1.1.3 设 $f: A \to B$,则

1) f 有左逆的充分必要条件为 f 是单射;

2) f 有右逆的充分必要条件为 f 是满射;

3) f 可逆的充分必要条件为 f 是双射.

证明 1) 设 f 有左逆 g,若 $f(x_1) = f(x_2)$,两边用 g 作用,得 $gf(x_1) = gf(x_2)$,即 $I_A(x_1) = I_A(x_2)$,得 $x_1 = x_2$,所以 f 是单射.

充分性:设 f 是单射,定义 B 到 A 的对应关系 g 为 $g(b) = a$,若 $b \in f(A)$,且 $f(a) = b$, $g(b) = a_1$,若 $b \in B - f(A)$,其中 a_1 是 A 中任意取定的一个元素.因 f 是单射,$g(b)$ 唯一确定,故 g 是映射.

又对于 $\forall a \in A$ 有 $gf(a) = g(b) = a$,所以 $g \cdot f = I_A$,g 是 f 的左逆.

2) 必要性:设 f 有右逆 h,则对于 $\forall b \in B$,有 $fh(b) = b$,即 $f(h(b)) = b$,即对 $\forall b \in B$,存在 $x = h(b)$ 使 $f(x) = b$,所以 f 是满射.

充分性:设 f 是满射,$\forall b \in B$,存在一个 a,使 $f(a) = b$,于是,我们定义一个 B 到 A 的对应关系 $h: b \to a$,则 h 是 B 到 A 的一个映射,且有 $fh(b) = f(h(b)) = f(a) =$

b,所以 $fh = I_B$,即 h 是 f 的右逆.

3）由 1),2) 直接可得.

关于逆映射有以下性质：

1）$(f^{-1})^{-1} = f$;

2）若 f 是 $A \to B$ 的可逆映射,g 是 $B \to C$ 的可逆映射,则 $g \cdot f$ 是 $A \to C$ 的可逆映射,且有$(gf)^{-1} = f^{-1}g^{-1}$.

当 A 是有限集时,A 上的一个变换 f 可逆的充分必要条件是 f 是单射,或只要 f 是 A 上的满射.

1.2　等价关系与集合的分类

利用映射可以建立起两个集合间的联系以利于对两个集合进行分析比较,有时我们还需要把一个集合分成若干个子集来进行研讨,这就需要有等价关系与集合的分类这样两个概念,本节将介绍这两个概念并进行一些初步的讨论.

1.2.1　二元关系

关系是一个使用非常广泛的基本概念,在日常生活中我们也熟悉这个词的含义.在数学上关系可以表达集合中元素间的联系,例如“3小于5”、“x大于y”、“点a在b与c之间”等.在给出关系这个概念的定义之前我们先分析两个例子.

例 1.2.1　某电影院的电影票与座位之间有“对号”关系.用 A 表示电影票的集合,用 B 表示座位的集合,用 R 表示“对号”关系,对于 $\forall a \in A$ 与 $b \in B$,则 a 与 b 有对号关系或没有对号关系两者必居其一,若 a 与 b 有对号关系,则说序对$(a,b) \in R$;若 a 与 b 没有对号关系,则说序对$(a,b) \notin R$,因此,全体有对号关系的序对作成一个集合就是 R,而 R 是卡氏积 $A \times B$ 的一个子集合.

例 1.2.2　考虑实数集 $\mathbf{R}$,$\forall a,b \in \mathbf{R}$,$a > b$ 与 $a \not> b$ 两者必居其一,令 A 表示实数的“$>$”关系.当 $a > b$ 时,就说$(a,b) \in \mathbf{R}$,当 $a \not> b$ 时,就说$(a,b) \notin \mathbf{R}$,因此,所有适合 $a > b$ 的序对(a,b) 作成的集合就是 $\mathbf{R}$,而 $\mathbf{R}$ 是卡氏积 $\mathbf{R} \times \mathbf{R}$ 的一个子集合.

定义 1.2.1　设 A,B 是任意两个集合,规则 R,对于 $\forall a \in A$ 与 $b \in B$,均可以确定 a 和 b 是否适合这个规则,若适合这个规则,就说 a 和 b 有**关系** R,记作 aRb,否则记作 $aR'b$.

A 和 B 之间的关系 R 也可以用 $A \times B$ 的一个子集来表示：$A \times B$ 的一个子集 R 称为 A 和 B 之间的关系.当$(a,b) \in R$ 时,说 a 和 b 有关系 R,记作 aRb;当$(a,b) \notin R$ 时,说 a 和 b 没有关系 R,记作 $aR'b$.

特别地,$A \times A$ 的一个子集 R 就称为 A 上的一个**二元关系**.

例 1.2.3 集合 $X=\{a,b\}$，$Y=\{c,d,e\}$，设 R 是集合 X 和 Y 间的一种规则：

$$aRc, aRd, aR'e, bR'c, bR'd, bRe.$$

它可以用 $X\times Y$ 的子集 $R=\{(a, c)、(a, d)、(b, e)\}$ 来表示.

例 1.2.4 在 $\mathbf{Z}$ 中取定一个正整数 n，规定

$$aRb \Leftrightarrow n \mid (a-b).$$

则 R 是 $\mathbf{Z}$ 上的一个二元关系，称为 $\mathbf{Z}$ 上的"模 n 同余关系"，记为 $a\equiv b(n)$. 用 $\mathbf{Z}\times\mathbf{Z}$ 的子集 $R=\{(a, b) \mid n \mid a-b\}$ 来表示.

1.2.2 等价关系

由以上例子可以看出，集合 A 上的二元关系有各种不同的类型，现在我们具体看看例 1.2.4 给出的 $\mathbf{Z}$ 上的模 n 同余关系的特性：

1）$\forall a\in\mathbf{Z}$，有 $n\mid a-a$，则 aRa，也就是说，$\forall a\in\mathbf{Z}$，都有$(a, a)\in R$；

2）如果 aRb，即 $n\mid a-b$，则有 $n\mid b-a$，则 bRa，如果$(a, b)\in R$，就有$(b, a)\in R$；

3）如果 aRb，bRc，即 $n\mid a-b$ 且 $n\mid b-c$，则有 $n\mid (a-b)+(b-c)$，即 $n\mid a-c$，也就是说，如果$(a, b)\in R$，$(b, c)\in R$，就有$(a, c)\in R$.

这样的三条性质并不是每个二元关系都能同时具有的，具有上述三条性质的二元关系在代数上特别重要，通常称为等价关系.

定义 1.2.2 设 R 是集合 A 上的二元关系，如果 R 具有下述性质：

1）反身性：aRa，对于 $\forall a\in A$；

2）对称性：如果 aRb，那么 bRa；

3）传递性：如果 aRb 且 bRc，那么 aRc.

那么称 R 是 A 上的一个**等价关系**. 通常用符号"$\sim$"表示等价关系. 当 $a\sim b$ 时，就说 a 与 b 是**等价**的.

例 1.2.5 集合 $M_n(\mathbf{R})$ 上的二元关系都是 $M_n(\mathbf{R})$ 上的等价关系.

$\sim_1=\{(A,B)\mid$ 存在可逆阵 P、Q，使得 $B=PAQ\}$；

$\sim_2=\{(A,B)\mid$ 存在可逆阵 P，使得 $B=P^{-1}AP\}$；

$\sim_3=\{(A,B)\mid$ 存在可逆阵 P，使得 $B=P'AP\}$.

1.2.3 集合的分类

为了进一步描述等价关系的性质，我们引进集合分类的概念.

定义 1.2.3 设 A 是任一非空集合，$A_i(i\in I)$ 为 A 的非空子集组成的以 I 为指标集的集合 $\sum=\{A_i\mid A_i\subseteq A, A_i\neq\varnothing, i\in I\}$，如果具有条件：

1) $\bigcup_{i\in I} A_i = A$;

2) $\forall i, j \in I$, 若 $i \neq j$,则 $A_i \cap A_j = \varnothing$.

那么称 $\{A_i \mid i \in I\}$ 是集合 A 的一个分类,每个 A_i 都称为 A 的一个类.

例 1.2.6　在整数集 $\mathbf{Z}$ 中,令 $A_r = \{n \in \mathbf{Z} \mid n = 4q + r\}$,$(r = 0,1,2,3)$,则 $\{A_r \mid 0,1,2,3\}$ 是 $\mathbf{Z}$ 的一个分类.

例 1.2.7　在二阶矩阵集合 $M_2(F)$ 中,令 $r_i = \{A \mid A \in M_2(F), 秩(A) = i\}$,$(i = 0,1,2)$,则 $\{r_0, r_1, r_2\}$ 是 $M_2(F)$ 的一个分类.

例 1.2.8　在实数集 $\mathbf{R}$ 中,令 $A_i = \{x \mid x \in \mathbf{R}, i-1 < x < i\}$,$(i \in \mathbf{Z})$,则 $\{A_i \mid i \in \mathbf{Z}\}$ 不是 $\mathbf{R}$ 的分类,因为每个整数 i 都不在任何 A_i 中,因此 $\bigcup_{i\in I} A_i \neq \mathbf{R}$.

同样令 $A_i = \{x \mid x \in \mathbf{R}, i-1 \leqslant x \leqslant i\}$,$(i \in \mathbf{Z})$,则 $\{A_i \mid i \in \mathbf{Z}\}$ 也不是 $\mathbf{R}$ 的分类,因为每个整数 i 既在 A_i 中又在 A_{i+1} 中,$A_i \cap A_{i+1} \neq \varnothing$.

1.2.4　等价关系与集合的分类

从以上例子可以看出,所谓集合的分类就是利用元素间的某种关系,把集合分解成一些称为类的子集. 比如在例 1.2.6 中,是利用"模 4 同余关系"给出了 Z 的一个分类,在例 1.2.7 中是利用"同秩矩阵关系"给出了 $M_2(F)$ 的一个分类. 那么集合 A 上的任一个二元关系是否一定能确定 A 的一个分类呢?事实上这是不可能的. 比如,在例 1.2.8 中给出了 $\mathbf{R}$ 上的两个二元关系,一个给的是"在同一开区间内"关系,另一个是把开区间改成闭区间,但它们都不能给出 $\mathbf{R}$ 的分类. 不难发现,例 1.2.6、例 1.2.7 中的二元关系都是等价关系,实际上等价关系的重要意义正是在于它是构成分类的一般准则.

定理 1.2.1　非空集合 A 中的任一等价关系都能确定 A 的一个分类;反之,A 的每一个分类也都能决定 A 的一个等价关系.

证明　设 $\sim$ 是 A 的一个等价关系,得到 A 的子集:$\forall a \in A$,所有与 a 等价的元素作成 A 的一个子集,记为 $\bar{a}$. 这样得到集合 $\sum = \{\bar{a} \mid a \in A\}$ 是 A 的一个分类.

1) 若 $a \sim b$,则 $\bar{a} = \bar{b}$;$\forall x \in \bar{a} \Leftrightarrow x \sim a \Leftrightarrow x \sim b \Leftrightarrow x \in \bar{b}$;

2) 若 $\bar{a} \neq \bar{b}$,则 $\bar{a} \cap \bar{b} = \varnothing$;

如果 $\bar{a} \cap \bar{b} \neq \varnothing$,则存在 $x \in \bar{a} \cap \bar{b}$ 使得 $x \sim a$,且 $x \sim b$,所以 $a \sim b$. 由 1) 知 $\bar{a} = \bar{b}$,与假设相矛盾,所以 $\bar{a} \cap \bar{b} = \varnothing$.

3) $\bigcup_{\bar{a}\in\Sigma} \bar{a} = A$;$\forall a \in A$,总有 $a \in \bar{a}$.

反之,设 $\sum = \{A_i \mid A_i \subseteq A, i \in I\}$ 是集合 A 的一个分类,利用这个分类给出 A 的一个等价关系"$\sim$",$a \sim b \Leftrightarrow a$ 与 b 在同一类.

下面证明关系 $\sim$ 是等价关系:

(1) $\forall a \in A$,显然有 $a \sim a$;

(2) 如果 $a \sim b$,即 a 与 b 在同一类,则 b 与 a 在同一类,即 $b \sim a$;

(3) 如果 $a \sim b$、$b \sim c$,即 a 与 b 在同一个类,b 与 c 在同一个类,因此 a 与 c 也必在同一个类,即 $a \sim c$,故 $\sim$ 是 A 的一个等价关系.

在定理中,$\sum = \{\bar{a} \mid a \in A\}$ 是 A 的一个分类,我们称子集$\bar{a}$ 是 A 的一个**等价类**,元素 a 称为类$\bar{a}$ 的**代表元**,由于 $\forall b \in \bar{a}$,均有$\bar{b} = \bar{a}$,这说明等价类$\bar{a}$ 中的每个元均可作为代表元,即等价类与其代表元的选取无关.

1.2.5 等价关系的商集

定义 1.2.4 设 $\sim$ 是 A 上的等价关系,由 $\sim$ 确定的一切等价类组成的集合,称为 A 关于等价关系 $\sim$ 的**商集**,记为 $A/\sim$.

例 1.2.9 设 $A = \{1, 2, 3, 4, 5\}$. 令 $A_1 = \{1, 2\}$,$A_2 = \{3\}$,$A_3 = \{4, 5\}$,则 $\sum = \{A_1, A_2, A_3\}$ 是 A 的一个分类,由分类 $\sum$ 所确定的等价关系 $\sim$ 是

$$\begin{aligned}\sim &= (A_1 \times A_1) \cup (A_2 \times A_2) \cup (A_3 \times A_3) \\ &= \{(1,1)、(2,2)、(1,2)、(2,1)\} \cup \{(3,3)\} \cup \{(4,4)、(5,5)、(4,5)、(5,4)\} \\ &= \{(1,1)、(2,2)、(3,3)、(4,4)、(5,5)、(1,2)、(2,1)、(4,5)、(5,4)\}.\end{aligned}$$

$\sum$ 也就是 A 关于 $\sim$ 的商集:

$$A/\sim = \{A_1, A_2, A_3\} = \{\bar{1}, \bar{2}, \bar{3}\}.$$

例 1.2.10 对 $\mathbf{Z}$ 利用模 n 同余关系 $\sim$,可获得 $\mathbf{Z}$ 的一个分类:

$$a \text{ 与 } b \text{ 同一个类} \Leftrightarrow a \sim b.$$

称为**模 n 剩余类**.

$\mathbf{Z}$ 关于模 n 同余关系的商集 $\mathbf{Z}/\sim$,通常用符号 $\mathbf{Z}_n$ 表示.

对 $\forall k \in \mathbf{Z}$,k 一定与 $0,1,2 \cdots n-1$ 这 n 个整数中的某一个整数是同余的,即它们被 n 除时的余数相同;其次,$0,1,2 \cdots n-1$ 中的任意两个数都是不同余的.

由这两点可知,$\mathbf{Z}$ 的模 n 剩余类,即 $\mathbf{Z}_n = \{\bar{0}, \bar{1}, \bar{2} \cdots \overline{n-1}\}$.

还要说明几点:

1) 若 $n = 0$,则规定 $\mathbf{Z}$ 中每个元都只与自己同余(模 0),所以 $\mathbf{Z}_0 = \{a \mid a \in \mathbf{Z}\}$,即 $\mathbf{Z}$ 的每个元都构成模 0 的一个剩余类,这时 $\mathbf{Z}_0$ 含有无限多个元.

若 $n \neq 0$,则 $\mathbf{Z}_n$ 中的每个元都含有无限多个整数,由于类中每个元都可以作为该类的代表,所以每个类的代表元有无限多个. 因此,今后在讨论 $\mathbf{Z}_n$ 的代数运算时,要证明

与代表元的选择无关.

2) 若 $n>0, m=-n$,则 $\mathbf{Z}_m=\mathbf{Z}_n$,事实上,

$$a\equiv b\ (\bmod\ n)\Leftrightarrow n\mid a-b\Leftrightarrow a-b=nq\Leftrightarrow a-b=m(-q)\Leftrightarrow a\equiv b(\bmod\ m).$$

今后如果没有特别声明,谈到模 n 剩余类集时,n 都指的是正整数.

3) 不能离开模 n 来谈剩余类.

在剩余类集 $\mathbf{Z}_2$ 中,元素 $\overline{0}=\{\cdots-6,-4,-2,0,2,4,6\cdots\}$.

但在 $\mathbf{Z}_3$ 中,元素 $\overline{0}=\{\cdots-9,-6,-3,0,3,6,9\cdots\}$.

可见,$\mathbf{Z}_2$ 中的剩余类 $\overline{0}$ 与 $\mathbf{Z}_3$ 中的剩余类 $\overline{0}$ 是不一样的.

1.3　偏序与全序

1.3.1　偏序

定义 1.3.1　设 S 是一个集合,$\leqslant$ 是 S 中一个二元关系,满足

(1) $\forall x\in S$,有 $x\leqslant x$;　　(反身性)

(2) $\forall x,y\in S$,若有 $x\leqslant y$ 且 $y\leqslant x$,则 $x=y$;　　(反对称性)

(3) $\forall x,y,z\in S$,若有 $x\leqslant y$ 且 $y\leqslant z$,则 $x\leqslant z$,.　　(传递性)

则称 $\leqslant$ 是 S 中一个**偏序**(partial ordering),S 称为**偏序集**,记作$(S,\leqslant)$.

若$(S,\leqslant)$还满足

(4) 对 $\forall x,y\in S$ 均有 $x\leqslant y$ 或 $y\leqslant x$,则称 $\leqslant$ 为 S 中一个**全序**(total ordering),$(S,\leqslant)$称为一个**全序集**.

全序集中任何两个元素均有序关系,而偏序集中则不一定.

注意到偏序集的子集仍是偏序集,如果 $x\leqslant y$ 且 $x\neq y$,则记为 $x<y$.

例 1.3.1　设 A 为任意集合,$S=2^A=\{X\mid X$ 是 A 的子集$\}$,在 S 中定义二元关系 $\leqslant$ 为:$X\leqslant Y\Leftrightarrow X\subseteq Y$,则不难检验 S 对 $\leqslant$ 满足定义 1.3.1 中条件(1)、(2)、(3),故$(S,\leqslant)$是偏序集,但不是全序集.

例 1.3.2　在正整数集合 $\mathbf{Z}^+$ 中定义 $\leqslant$ 为整除关系,即 $a\leqslant b\Leftrightarrow a\mid b$,则$(2^A,\mid)$是偏序集,但不是全序集. 如果我们在 $\mathbf{Z}^+$ 中定义 $\leqslant$ 就是普通的小于或等于关系,则$(2^A,\leqslant)$是全序集.

例 1.3.3　域上的 n 维向量空间 V 的所有子空间对于包含关系"$\subset$"构成一个偏序集,但不是全序集.

下面给出偏序集中最(极)大(小)元及子集的上(下)界的概念.

定义 1.3.2　1) 设 $a\in S$,若 $\forall x\in S$ 有 $x\leqslant a\ (x\geqslant a)$,则称 a 是 S 的**最大(小)元**.

2) 设 $a \in S$,若 $x \geqslant a$ $(x \leqslant a)$,则 $x = a$,则称 a 是 S 中的一个**极大(小)元**.

3) 设 T 是 S 的一个子集,$a \in S$,若对 $\forall x \in T$ 均有 $x \leqslant a$ $(x \geqslant a)$,就称 a 是 T 的一个**上(下)界**. 注意:子集的上(下)界未必在此子集中.

4) 设 $T \subseteq S$,a 是 T 的一个上界,若对于 T 的任意一个上界 a' 均有 $a \leqslant a'$,则称 a 是 T 的**最小上界**. 类似有**最大下界**的概念.

例 1.3.4 $\mathbf{Z}^+ = \{1,2,3,\cdots\}$ 对整除关系构成一个偏序集;

设 $S = \{1,2,3,4,5,6\}$,S 有最小元 1,无最大元,4,5,6 都是 S 的极大元. S 在 $\mathbf{Z}^+$ 中的上界有很多,4,5,6 的公倍数都是,但最小上界只有一个,即 4,5,6 的最小公倍数 60. 这个上界不在 S 中.

1.3.2 全序集的良序性

定义 1.3.3 设 A 为全序集,若 A 的任何非空子集有最小元,则称 A 是**良序集**.

若集合 A 的元素间的关系 $\sim$ 使 A 成为一个良序集,则称 $\sim$ 是 A 的一个**良序关系**.

数集 $\mathbf{Z}^+$ 是良序集:设 M 是 $\mathbf{Z}^+$ 的任一非空子集,可在 M 中任取一数,设为 n,则 M 中小于或等于 n 的数只有有限个(不多于 n 个),故存在一个最小数. 所以 $\mathbf{Z}^+$ 是良序集.

整数集合 $\mathbf{Z}$ 对数的大小不是良序的,但可对 $\mathbf{Z}$ 重新规定序使其成为良序集.

由自然数集的良序性可得以下的数学归纳法原理.

定理 1.3.1 设 M 是由自然数构成的集合,若 $1 \in M$,而且当 $n-1 \in M$ 时必有 $n \in M$,则 M 是自然数集.

证明 设 $N = \mathbf{Z}^+ - M$,若 $N \neq \varnothing$,则由 $\mathbf{Z}^+$ 的良序性知 N 有最小数 a,且因 a 不属于 M 知 $a-1$ 不属于 N,故 $a-1 \in \mathbf{Z}^+$. 由 a 在 N 中的极小性知 $a-1$ 不属于 N,于是 $a-1 \in M$,由定理所给条件得 $a \in M$,矛盾. 所以 $N = \varnothing$,即 $M = \mathbf{Z}^+$.

如果一个命题与自然数有关,有以下的数学归纳法:

首先证明命题对 1 成立,然后假设命题对 $n-1$ 成立,若能证明命题对 n 也成立,则此命题对所有自然数都成立.

数学归纳法还有第二种形式:首先证明命题对 1 成立,然后假设命题对所有小于 n 的自然数都成立,若能证明命题对 n 也成立,则命题对所有自然数都成立.

1.3.3 超限归纳法

还可把数学归纳法推广到任何良序集,这就是所谓的**超限归纳法**.

定理 1.3.2(超限归纳法原理) 设 $(S, \leqslant)$ 是一个良序集,$P(x)$ 是与元素 $x \in S$ 有关的一个命题,如果

1) 对于 S 中的最小元 a_0,$P(a_0)$ 成立.

2) 假定对 $\forall x < a$,$P(x)$ 成立,可证明 $P(a)$ 也成立,则 $P(x)$ 对 $\forall x \in S$ 都成立.

近代数学的严谨性在于它的公理化，而当代数学的许多重要结论的公理基础之一是 Zermelo 公理(选择公理)，但在许多理论研究中人们更喜欢应用与选择公理等价的 Zorn 引理.

有时为清晰起见，把若干个集合作成的集合称为类. 在集合的包含关系下，每个类就是一个偏序集.

定理 1.3.3(逆归定理) 假设 S 是一个集合，$a\in S$，并且对 $\forall n\in\mathbf{N}$，$f_n: S\to S$ 均是函数，则存在唯一的函数 $\varphi: \mathbf{N}\to S$，使得 $\varphi(0)=a$，并且 $\varphi(n+1)=f_n(\varphi(n))$，($\forall n\in\mathbf{N}$).

证明 构造 $\mathbf{N}\times S$ 上的一个关系 R，使得它是满足上述性质的函数 $\varphi: \mathbf{N}\to S$ 的图象，令

$\psi=\{Y\subset\mathbf{N}\times S\mid(0,a)\in Y$，并且 $(n, x)\in Y$，则 $(n+1,f_n(x))\in Y$，$(\forall n\in\mathbf{N})\}$，

由于 $\mathbf{N}\times S\in\psi$，从而 $\psi\neq\varphi$. 令 $R=\bigcap Y$，则 $R\in\psi$. 又设 M 为子集合

$$M=\{n\in\mathbf{N}\mid \text{存在唯一的 } x_n\in S\text{，使得}(n, x_n)\in R\}.$$

我们用归纳法证明 $M=\mathbf{N}$. 如果 $0\in M$，则有 $(0,b)\in R$，其中 $b\neq a$，并且集合 $R-\{(0,b)\}\subset\mathbf{N}$，从而 $R=\bigcap Y\subset R-\{(0,b)\}$，这就导致矛盾. 因此 $0\in M$.

现在假定 $n\in M$(即有唯一的 $x_n\in S$ 使得 $(n, x_n)\in R$)，则 $(n+1, f_n(x_n))\in R$. 如果又有 $(n+1, c)\in R$，而 $c\neq f_n(x_n)$，则 $R-\{(n+1, c)\}\in\psi$，由此又可像上面那样导致矛盾. 因此 $x_{n+1}=f_n(x_n)$ 是 S 中唯一的元素，使得 $(n+1, x_{n+1})\in R$，于是由归纳法可知 $\mathbf{N}=M$，即 $n\to x_n$，定义了一个函数 $g: \mathbf{N}\to S$，它的图象为 R，由于 $(0, a)\in R$，从而 $g(0)=a$. 对于每个 $n\in\mathbf{N}$，$(n, x_n)=(n, g(n))\in R$，由于 $R\in\psi$，从而 $(n+1, f_n(g(n)))\in R$. 但是 $(n+1, x_{n+1})\in R$，由 x_{n+1} 的唯一性推出

$$g(n+1)=x_{n+1}=f_n(g(n)).$$

如果 A 是非空集合，A 中的序列是一个函数 $\mathbf{N}\to A$，一个序列通常表示成 $\{a_0, a_1\cdots\}$，或者 $\{a_n\}$，其中 $a_n\in A$ 是 $n\in\mathbf{N}$ 的象.

类似地，函数 $\mathbf{N}^*\to A$ 也称作序列，并且表示成 $\{a_0,a_1\cdots\}$，或者 $\{a_n\}$ $n\in\mathbf{N}^*$，这些符号不会引起混乱.

1.3.4 Zorn 引理及等价命题

定理 1.3.4(选择公理) 对若干个非空集合作成的每个类 F 必存在一个函数 f(称为 F 上的一个选择函数)，使得对每一个 $A\in F$，我们有 $f(A)\in A$.

定理 1.3.5(Zorn 引理) 如果偏序集 $(S, \leqslant)$ 的任何全序子集均在 S 中有上界，则 S 必含有极大元.

关于选择公理与 Zorn 引理的等价性证明一般篇幅很长，可参见：谢邦杰《超穷书与超穷论法》，吉林人民出版社 1979 年版.

与以上两个公理等价的另一个命题是：

定理 1.3.6(良序定理) 在任何集合上存在良序关系，即任何集合可以排成一个良序集.

作为 Zorn 引理的一个应用，我们给出由 Zorn 引理推出良序定理的证明，以便读者体会如何应用 Zorn 引理.

证明 设 A 是任一集合，则 A 的每一个子集上的良序关系都是 $A\times A$ 的一个子集. 以 Φ 来表示 A 的每个子集上的良序关系的族 Φ，作为 $A\times A$ 的子集族的类 Φ，可以根据包含关系定义偏序关系如下：设 $\sigma,\tau\in\Phi$，$\sigma\leqslant\tau\Leftrightarrow\sigma\subset\tau$，

对于偏序集 Φ，由于 F 的任一全序子集还是一个个依次包含的 $A\times A$ 的一族子集，这族子集的并便是这族子集的上界. 应用 Zorn 引理，F 中应有一个极大元 ρ，显然 ρ 是 $A\times A$ 的一个子集.

假设 ρ 仅是 A 的真子集 B 的序关系，设 $y\in A-B$，则 $\varphi=\rho\cup\{(x,\ y)\mid x\in B\}$ 是 A 的子集 $B\cup\{y\}$ 上的序关系，而且 $\rho\subset\varphi$，$\rho\neq\varphi$，这与 ρ 的极大性矛盾，得证.

1.4 基 数

1.4.1 基数(势)的基本性质

我们说两个集合 S 和 T 有相同的基数性，记作 $|S|=|T|$，如果这两个集合之间存在一个双射函数(一一对应)，读者可能会想到：$|\mathbf{Z}|=|\mathbf{N}|$，$|\mathbf{Q}|=|\mathbf{N}|$，其中 $\mathbf{N}$，$\mathbf{Z}$ 和 $\mathbf{Q}$ 分别为自然数、整数和有理数.

如果 S 与 T 的一个子集一一对应，我们就记作 $|S|\leqslant|T|$. 如果 S 与 T 的一个真子集一一对应，且 $|S|\neq|T|$，我们就记作 $|S|<|T|$. 第二个条件是必要的，因为例如，$\mathbf{N}$ 与 $\mathbf{Z}$ 的一个真子集一一对应，但 $\mathbf{N}\not<\mathbf{Z}$.

我们这儿就不详细讨论基数了，而是把注意力集中在集合的基数性上，它表示了集合的“大小”. 比起要清楚地说明一给定集合的大小(基数性)来，讨论两个集合有相同或不同的大小(基数性)要容易得多.

我们给每个集合 S 一个基数，记为 $|S|$ 或 card(S)，用来测量集合的大小. 实际上，基数恰好是集合的非常特殊的类型. 但是，我们只能简单地把它们想成是用来测量集合大小的模糊的对象.

一个集合是有限的，如果对于任意正整数 n，它能在形为 $\mathbf{Z}_n=\{0,\ 1,\ \cdots,\ n-1\}$ 的一个集合的一一对应中. 一个有限集的基数(或基数性)正好是这一集合中元素的个数. 自然数集 $\mathbf{N}$ 的基数是 $\aleph_0$. 因此 $|\mathbf{N}|=|\mathbf{Z}|=|\mathbf{Q}|=\aleph_0$.

$\aleph_0$ 的任意集合叫做**可数无限集**. 而任何有限集或可数无限集叫做**可数集**. 因为能

够证出 $|\mathbf{R}|>|\mathbf{N}|$,所以实数是不可数的.

如果 S 和 T 是有限集,那么 $|S|\leqslant|T|$ 且 $|T|\leqslant|S|\Rightarrow|S|=|T|$.

1.4.2 基数的比较

定理 1.4.1 在全体集合所构成的类 T 上,等势是一个等价关系.

证明 注意 $\Phi\sim\Phi$,由于 $\Phi\subset\Phi\times\Phi$ 是一个关系,它(无条件地)为一一对应函数;假设 $I_0=\Phi$,并且对于每个 $n\in\mathbf{N}^*$,令 $I_n=\{1,2,3\cdots n\}$.

不难证明 I_m 和 I_n 等势 $\Leftrightarrow m=n$. 我们说一个集合 A 恰好有 n 个元素,这意味着 A 和 I_n 等势. 这样的集合(即对某个(唯一的)$n\geqslant 0$,$A\sim I_n$)叫作有限集合,否则便叫作无限集合. 于是,对于有限集合 A,A 的等势等价类给出下面问题的答案:A 中包含多少元素?基于这些考虑我们给出:

定义 1.4.1 集合 A 的**基数(势)**是 A 对于等势关系的等价类,表示成 $|A|$,$|A|$ 称作无限或有限势,视 A 为无限集合或者有限集合而定.

势也用 α,β,γ 等小写希腊字母表示,基于上面一段所指出的原因,我们把整数 $n\geqslant 0$ 等同于势 $|I_n|$,并且写成 $|I_n|=n$,从而有限集合的势恰好是集合中的元素个数.

通常定义势的方法与我们上面所作的稍有不同,通常的定义可表明势实际上是一个集合(不像定义 1.4.1 中那样只是一个本性类).

1) 每个集合都有唯一的势;

2) 两个集合有同样的势,当且仅当它们是等势的(即 $|A|=|B|\Leftrightarrow A\sim B$);

3) 有限集合的势是该集合中的元素个数.

于是,关于势的命题只不过是关于集合等势性的命题.

例 1.4.1 自然数集 $\mathbf{N}$ 的势习惯上表为 $\aleph_0$. 势为 $\aleph_0$ 的集合 A 叫作可数集合. 非零正整数集合 $\mathbf{N}^*$,整数集 $\mathbf{Z}$ 和有理数集 $\mathbf{Q}$ 均是可数集合,但是实数集 $\mathbf{R}$ 不是可数的.

定义 1.4.2 假设 α 和 β 为势,A 和 B 是非交集合,并且 $|A|=\alpha$,$|B|=\beta$,则和 $\alpha+\beta$ 定义为势 $|A\cup B|$,而积 $\alpha\beta$ 定义为势 $|A\times B|$.

在定义积 $\alpha\beta$ 的时候,实际上不必假定 A 和 B 是非交的. 从势 α 的定义可知,永远存在集合 A,使得 $|A|=\alpha$. 不难证明,定义 $\alpha+\beta$ 时所需要的非交集合是永远存在的,并且和 $\alpha+\beta$ 及积 $\alpha\beta$ 均与集合 A 和 B 的选取方式无关. 势的加法和乘法满足结合律和交换律,并且分配律也成立. 此外,有限势的加法和乘法与它们所等同的非负整数之间的加法和乘法是一致的,这是因为,如果 A 和 B 分别有 m 和 n 个元素,并且 $A\cap B=\varnothing$,则 $A\cup B$ 有 $m+n$ 个元素,而 $A\times B$ 有 mn 个元素.

定义 1.4.3 假设 α 和 β 为势,A 和 B 为集合,并且,$|A|=\alpha$,$|B|=\beta$,如果 A 与 B 的一个子集合等势(即存在一个单射 $A\rightarrow B$),我们便称 α 小于或等于 β(表示成 $\alpha\leqslant\beta$ 或者 $\beta\geqslant\alpha$). 如果 $\alpha\leqslant\beta$ 并且 $\alpha\neq\beta$,则称 α 严格小于 β(表示成 $\alpha<\beta$ 或

者$\beta > \alpha$).

1.4.3 施罗德-伯恩斯坦定理

定理 1.4.2 如果 A 是集合而 $P(A)$ 是它的幂集合,则 $|A| < |P(A)|$.

证明 $a \to \{a\}$,$A \to P(A)$ 是一个单射,从而 $|A| \leqslant |P(A)|$. 如果存在一一对应 f: $A \to P(A)$,那么对于集合 $B = \{a \in A \mid a \notin f(a)\} \subset A$,存在着某个 $a_0 \in A$,使得 $f(a_0) = B$. 但这就导致矛盾: $a_0 \in B$ 同时又 $a \notin B$,因此 $|A| \neq |P(A)|$,即 $|A| < |P(A)|$.

定理 1.4.3

1) **(施罗德-伯恩斯坦定理)** 对于任意集合 S 和 T, $|S| \leqslant |T|$ 且 $|T| \leqslant |S| \Rightarrow |S| = |T|$.

2) **(康托尔定理)** 如果 $\Pi(S)$ 表示 S 的幂集,那么 $|S| < |\Pi(S)|$.

3) 若 S 是一个无限集,那么 $|S| = |\Pi_0(S)|$.

证明 1) 我们接着哈尔莫斯(1960)的证明. 令 f: $S \to T$ 是 S 到 T 的一个单射函数,g: $T \to S$ 是 T 到 S 的一个单射函数. 我们要说明存在一个 S 到 T 的双射函数. 我们构造以下定义: 元素 $s \in S$ 有**后项** $f(s)$, $g(f(s))$, $f(g(f(s)))$, …

如果 t 是 s 的后项,那么 s 是 t 的前项,类似地,我们定义 t 的后项和 s 的前项. 现在,通过找出一个元素的首项,我们发现存在三种可能 —— 这一元素的首项可能在 S 中,或 T 中,或没有首项. 相应地,可以把 S 写成以下三种不相交集:

$$S_S = \{s \in S \mid s \text{ 的首项在 } S \text{ 中}\},$$

$$S_T = \{s \in S \mid s \text{ 的首项在 } T \text{ 中}\},$$

$$S_\infty = \{s \in S \mid s \text{ 没有首项}\} \text{ 的并}.$$

同样,我们把 T 写成 T_S,T_T 和 T_∞ 的不相交的并.

限制 $f|_{S_S}$: $S_S \to T_S$ 是一个双射. 因为如果 $t \in T_S$,那么对于 $\forall s' \in S$, $t = f(s')$. 但是 s' 和 t 的首项相同,所以 $s' \in S_S$. 对于映射 $g|_{T_T}$: $T_T \to S_T$ 和 $f|_{S_\infty}$: $S_\infty \to T_\infty$ 也是双射. 把这三个双射加起来,就得到 S 与 T 之间的一个双射,从而 $|S| = |T|$.

2) 由 $\varepsilon(s) = \{s\}$ 定义的包含映射 ε: $S \to P(S)$ 是 S 到 $P(S)$ 的一个单射,所以 $|S| \leqslant |P(S)|$. 为了完成康托尔定理的证明,我们必须说明如果 f: $S \to P(S)$ 是任意单射,那么 f 不是满射的. 为了这一目标,令 $X = \{s \in S \mid s \notin f(s)\}$,那么 $X \in P(S)$.

现在我们证明 X 不在 $\mathrm{Im}(f)$ 中. 对于 $\forall x \in X$,设 $X = f(x)$,那么如果 $x \in X$,根据 X 的定义,有 $x \notin X$. 另一方面,如果 $x \notin X$,也根据 X 的定义,有 $x \in X$. 这一矛盾表明 $X \notin \mathrm{Im}(f)$. 所以 f 不是满射.

3) 证明略.

1.4.4 基数的运算

现在我们定义基数的加法、乘法和乘方.如果S和T是集合,那么笛卡儿积$S\times T$是全体有序对所成集合$S\times T=\{(s,t)\mid s\in S, t\in T\}$.

同样,令S^{T}表示T到S的所有映射所组成的集合.

定理 1.4.4 令κ,λ和μ是基数,那么以下性质成立:

1)**(结合性)** $\kappa+(\lambda+\mu)=(\kappa+\lambda)+\mu$且$\kappa(\lambda\mu)=(\kappa\lambda)\mu$.

2)**(交换性)** $\kappa+\lambda=\lambda+\kappa$且$\kappa\lambda=\lambda\kappa$.

3)**(分配性)** $\kappa(\lambda+\mu)=\kappa\lambda+\kappa\mu$.

4)**(指数的性质)** a) $\kappa^{\lambda+\mu}=\kappa^{\lambda}\kappa^{\mu}$. b) $(\kappa^{\lambda})^{\mu}=\kappa^{\lambda\mu}$. c) $(\kappa\lambda)^{\mu}=\kappa^{\mu}\lambda^{\mu}$.

定理 1.4.5 令κ,λ是基数,其中至少有一个是无限的,那么

1) $\kappa+\lambda=\max\{\kappa,\lambda\}$. 2) $\kappa\lambda=\max\{\kappa,\lambda\}$.

定理 1.4.6 1) 对基数$\aleph_0$运用有限次加法和乘法后,结果仍然为$\aleph_0$.

特别地,对任意非零$n\in\mathbf{N}$,$n\cdot\aleph_0=\aleph_0$,$\aleph_0^n=\aleph_0$.

2) 对基数$2^{\aleph_0}$运用可数次加法和乘法后,结果仍为$2^{\aleph_0}$.

特别地

$$\aleph_0\cdot 2^{\aleph_0}=2^{\aleph_0},(2^{\aleph_0})^{\aleph_0}=2^{\aleph_0}.$$

运用这一定理,我们可以建立其他关系,比如

$$2^{\aleph_0}\leqslant(\aleph_0)^{\aleph_0}\leqslant(2^{\aleph_0})^{\aleph_0}=2^{\aleph_0}.$$

根据施罗德-伯恩斯坦定理,有$(\aleph_0)^{\aleph_0}=2^{\aleph_0}$.

定理 1.4.7 用集合K表示集族$\{A_k\mid k\in K\}$,且$|K|=\kappa$.如果对于$\forall k\in K$,$|A_k|\leqslant\lambda$,那么

$$\left|\bigcup_{k\in K}A_k\right|\leqslant\lambda\kappa.$$

定理 1.4.8 每个无限集均有可数子集,特别地,对于无限基数α,均有$\aleph_0\leqslant\alpha$.

证明 如果B是无限集合A的有限子集合,则$A\rightarrow B$非空,对于A的每个有限子集合B,取一个元素$x_B\in A-B$(选择公理).以Φ表示A的全部有限子集合所构成的集合,定义映射$f:\Phi\rightarrow f(B)\cup\{x_B\}$.取$a\in A$.由递归定理(对于每个$n$均取$f_n=f$)可知存在函数$g:\mathbf{N}\rightarrow\Phi$使得$g(0)=\{a\}$,并且

$$g(n+1)=f(g(n))=g(n)\cup\{x_{g(n)}\}\ (n\geqslant 0).$$

设函数$h:\mathbf{N}\rightarrow A$定义为

$$h(0)=a,h(1)=x_{g(0)}=x_{\{a\}},\cdots,h(n+1)=x_{g(n)}.$$

利用 $\mathbf{N}$ 的序性质及以下诸事实可以证明 h 为单射：

1) $h(n) \in g(n)$ $(\forall n \geqslant 0)$；

2) $h(n) \notin g(n-1)$ $(\forall n \geqslant 1)$；

3) $h(n) \notin g(m)$ $(\forall m < n)$.

从而 $\mathrm{Im}h$ 是 A 的子集合，并且 $|\mathrm{Im}h| = |\mathbf{N}| = \xi_0$.

引理 1.4.9　如果 A 是无限集合，F 是有限集合，则 $|A \cup F| = |A|$，特别地，对于每个无限势 α 和每个自然数(有限基数)n，$\alpha + n = \alpha$.

证明　必要时用 $F-A$ 代替 F，不妨设 $A \cap F = \varnothing$. 如果 $F = \{b_1, b_2 \cdots b_n\}$ 而 $D = \{X_i \mid i \in \mathbf{N}^*\}$ 是 A 的一个可数子集，证 $f: A \to A \cup F$ 是一一对应，其中 f 定义为

如果 $x = x_i (1 \leqslant i \leqslant n)$，$f(x) = b_i$；如果 $x = x_i (i > n)$，$f(x) = x_{i-n}$；如果 $x \in A - D$，$f(x) = x$.

定理 1.4.10　如果 α 和 β 为势，$\beta \leqslant \alpha$ 并且 α 为无限基数，则 $\alpha + \beta = \alpha$.

定理 1.4.11　如果 α 和 β 为基数，$0 \neq \beta \leqslant \alpha$ 并且 α 是无限的，则 $\alpha\beta = \alpha$. 特别地，$\alpha\xi_0 = \alpha$. 又若 β 是有限的，则 $\xi_0\beta = \xi_0$.

证明　由于 $\alpha \leqslant \alpha\beta \leqslant \alpha\alpha$，从而(像定理 1.4.6 的证明那样)只需证明 $\alpha\alpha = \alpha$. 假设 A 是无限集合，$|A| = \alpha$，而令 $T = \{$一一对应 $f: X \times X \to X \mid X$ 为 A 的无限子集合$\}$，由于 A 有可数子集合 D (从而 $|D| = |\mathbf{N}| = |\mathbf{N}^*|$)，并且映射 $\mathbf{N}^* R \mathbf{N}^* \to \mathbf{N}^*$，$(m, n) \to t2^{m-1})(2n-1)$ 是一一对应，可知 $T \neq \varnothing$，将 T 赋以扩充序，利用 Zorn 引理得到极大元 $g: B \times B \to B$. 从 g 的定义可知 $|B \times B| = |B|$. 从而为了完成证明，只需证 $|B| = |A| = \alpha$. 假如 $|A—B| > |B|$，有 $A—B$ 的子集 C，使得 $|C| = |B|$. 证明 $|C| = |B| = |B \times B| = |B \times C| = |C \times B| = |C \times C|$，并且这些集合是彼此非交的，于是

$$
\begin{aligned}
|(B \cup C) \times (B \cup C)| &= |(B \times B) \cup (B \times C) \cup (C \times B) \cup (C \cup C)| \\
&= |B \times B| + |B \times C| + |C \times B| + |C \times C| \\
&= (|B| + |B|) + (|C| + |C|) \\
&= |B| + |C| = |B \cup C|.
\end{aligned}
$$

从而有一一对应 $(B \cup C) \times (B \cup C) \to (B \cup C)$，这与 T 中 g 的极大性矛盾，因此 $|A - B| \leqslant \beta$，并且 $|B| = |A - B| + |B| = |(A - B) \cup B| = |A| = \alpha$.

定理 1.4.12　假设 A 为集合，并且对于每个整数 $n \geqslant 1$，令 $A^n = A \times A \times \cdots \times A$.

1) 如果 A 是有限的，则 $|A^n| = |A|^n$，如果 A 是无限的，则 $|A^n| = |A|$；

2) $|\bigcup\limits_{n \in \mathbf{N}^*} A^n| = \aleph_0 |A|$.

证明　1) 如果 $|A|$ 有限，则命题显然成立. 如果 $|A|$ 无限，可以对 n 归纳证明 (对于 $n = 2$ 的情形即是定理 1.4.11).

2）集合 $A^n(n \geqslant 1)$ 是彼此非交的，如果 A 是无限的，由 1）可知对每个 n 均有一一对应

$$f_n: A^n \to A, \bigcup_{n \in \mathbf{N}^*} A^n \to \mathbf{N}^* \times A, u\ (u \in A^n) \mapsto (n, f_n(u))$$

是一一对应，因此

$$\left| \bigcup_{n \in \mathbf{N}^*} A^n \right| = |\mathbf{N}^* \times A| = |\mathbf{N}^*||A| = \xi_0 |A|.$$

如果 A 是空集，则 2）显然成立，于是设 A 为有限非空集合. 这时每个 A^n 都是非空的，并且易证 $|\mathbf{N}^*| \leqslant \left| \bigcup_{n \in \mathbf{N}^*} A^n \right|$. 进而每个 A^n 均是有限的，从而对每个 n 均有单射 g_n: $A^n \to \mathbf{N}^*$，映射

$$\bigcup_{n \in \mathbf{N}^*} A^n \to \mathbf{N}^* \times \mathbf{N}^*, u\ (\in A^n) \mapsto (n, g_n(u))$$

是单射，从而由定理 1.4.7 有

$$\left| \bigcup_{n \in \mathbf{N}^*} A^n \right| \leqslant |\mathbf{N}^* \times \mathbf{N}^*| = |\mathbf{N}^*| = \xi_0.$$

于是由 Berrlsteh 定理可知 $\left| \bigcup_{n \in \mathbf{N}^*} A^n \right| = \xi_0$. 但是由于 A 为有限集合，从而 $\xi_0 = \xi_0 |A|$.

推论 1.4.13　如果 A 为无限集合，而 $F(A)$ 为 A 的所有有限子集构成的集合，则 $|F(A)| = |A|$.

第2章 多项式与矩阵代数理论

多项式是线性代数的重要组成部分,也是代数学中的一个基本概念,并且是中学数学中的主要讨论对象之一.多项式代数以多项式因式分解存在唯一性定理为中心展开整个理论体系,主要讨论数域 F 上的整除性理论.注意在数域 F 与数环 $\mathbf{R}$ 上讨论多项式理论的区别.

2.1 一元多项式理论

2.1.1 一元多项式定义及运算

定义 2.1.1 设 $\mathbf{R}$ 为含有数 1 的数环,形式表达式

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n \tag{2.1}$$

称为数环 $\mathbf{R}$ 上关于 x 的多项式,简称**一元多项式**,其中 n 为非负整数,$a_0, a_1, \cdots, a_n \in \mathbf{R}$,通常用符号 $f(x)$ 表示.

多项式(2.1)中 $a_i x^i$ 称为 i 次项,a_i 称为 i 次项的系数,a_0 称为常数项.

加法:$\forall f(x) = \sum_{i=1}^{n} a_i x^i, g(x) = \sum_{i=1}^{n} b_i x^i$,则 $f(x)+g(x) = \sum_{i=1}^{n}(a_i+b_i)x^i$;

乘法:$\forall f(x) = \sum_{i=1}^{n} a_i x^i, g(x) = \sum_{j=1}^{m} b_j x^j$,则 $f(x)g(x) = \sum_{k=1}^{n+m}(\sum_{i+j=k} a_i b_i)x^k$;

满足下列算律

(1) $f(x)+g(x) = g(x)+f(x)$;

(2) $(f(x)+g(x))+h(x) = f(x)+(g(x)+h(x))$;

(3) $f(x)+(-f(x)) = 0$;

(4) $f(x)g(x) = g(x)f(x)$;

(5) $(f(x)g(x))h(x) = f(x)(g(x)h(x))$;

(6) $(f(x)+g(x))h(x) = f(x)h(x)+g(x)h(x)$;

(7) 若 $f(x), g(x) \in \mathbf{R}[x]$,且不全为零,若 $f(x)h(x) = g(x)h(x)$,且 $h(x) \neq 0$,则 $f(x) = g(x)$.

以上(2) ～ (7) 的定义都是在数域 F 上一元多项式环 $F[x]$ 中给出的.

集合 $\mathbf{R}[x] = \{a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \mid a_i \in \mathbf{R}, i = 0,1 \cdots n, \forall n(\geqslant 0) \in \mathbf{Z}\}$ 对多项式的加法、乘法构成的环称为数环 $\mathbf{R}$ 上的**一元多项式环**.

例 2.1.1　求满足条件的非零复系数多项式 $f(x)$，$f(f(x)) = (f(x))^n$，其中 n 为正整数.

解　1) 若 $\partial^0 f(x) = 0$，则 $f(x) = c$，其中常数 $c \neq 0$. 于是由题设 $c = f(f(x)) = (f(x))^n = c^n$，即 $c^{n-1} = 1$，其中解为 $n-1$ 次单位根 $c_k = \cos\dfrac{2k\pi}{n-1} + \mathrm{i}\sin\dfrac{2k\pi}{n-1}$，$k = 0$，$1 \cdots n-2$，这样 $f(x) = c_k, k = 0, 1 \cdots n-2$，且 $(c_k)^{n-1} = 1$，即为所求.

2) 若 $\partial^0 f(x) > 0$，设 $\partial^0 f(x) = t$，则由题设比较等式两端多项式的首项次数得 $t^2 = tn$，即 $t = n$，于是 $f(x)$ 为 n 次多项式，令 $f(x) = a_0x^n + \cdots a_{n-1}x + a_n$，则 $f(f(x)) = a_0(f(x))^n + \cdots + a_{n-1}f(x) + a_n = (f(x))^n$，$(a_0 - 1)(f(x))^n + \cdots + a_{n-1}f(x) + a_n = 0$，$a_0 = 1, a_1 = \cdots = a_{n-1} = a_n = 0$，故 $f(x) = x^n$.

定义 2.1.2　$\forall f(x) \in \mathbf{R}[x], f(x) \neq 0$，当 $a_n \neq 0$ 时，a_nx^n 称为多项式 $f(x)$ 的**首项**，a_n 为**首项系数**，n 称为 $f(x)$ 的**次数**，用符号 $\partial^0 f(x)$ 表示，即 $\partial^0 f(x) = n$.

特别地，系数全为零的多项式称为零多项式，记为 0. 零多项式不定义次数.

定理 2.1.1(次数定理)　设 $f(x), g(x) \in \mathbf{R}[x]$，且 $f(x) \neq 0, g(x) \neq 0$，则

1) 当 $f(x) + g(x) \neq 0$ 时，$\partial^0(f(x) + g(x)) \leqslant \max\{\partial^0 f(x), \partial^0 g(x)\}$；

2) $\partial^0(f(x)g(x)) = \partial^0 f(x) + \partial^0 g(x)$.

2.1.2　多项式的整除

定义 2.1.3　设 $f(x), g(x) \in F[x]$，若存在 $h(x) \in F[x]$，使 $f(x) = g(x)h(x)$，则称 $g(x)$ 整除 $f(x)$，记为 $g(x) \mid f(x)$，其中 $g(x)$ 称为 $f(x)$ 的一个因式.

多项式理论是围绕多项式的整除而展开的，首先考虑多项式的整除性质：

(1) 若 $f(x) \mid g(x)$，$g(x) \mid h(x)$，则 $f(x) \mid h(x)$；

(2) 若 $f(x) \mid g(x)$，$f(x) \mid h(x)$，则 $f(x) \mid g(x) + h(x)$；

(3) 若 $f(x) \mid h(x)$，对于 $\forall g(x) \in F[x]$，有 $f(x) \mid g(x)h(x)$；

(4) 若 $f(x) \mid h_i(x)$，对 $\forall g_i(x) \in \mathbf{R}[x]$，有 $f(x) \mid \sum\limits_{i=1}^{n} g_i(x)h_i(x)$，其中 $i = 1$，$2 \cdots n$；

(5) $\forall c(\neq 0) \in F, \forall f(x) \in F[x]$，有 $c \mid f(x), cf(x) \mid f(x)$；

(6) 若 $f(x) \mid g(x)$，$g(x) \mid f(x)$，则 $g(x) = cf(x)$，其中 $c(\neq 0) \in F$.

例 2.1.2　求证：$x^d - 1 \mid x^n - 1 \Leftrightarrow d \mid n$.

证明　充分性　由 $d \mid n$，得 $n = dq$，则 $x^n - 1 = (x^d)^q - 1$，故 $x^d - 1 \mid x^n - 1$.

必要性　假设 $n = dq + r$,其中 $r = 0$ 或 $r < d$,于是

$$x^n - 1 = (x^d)^q x^r - 1 = x^n - 1 = x^r[(x^d)^q - 1] + (x^r - 1),$$

由 $x^d - 1 \mid x^n - 1$ 推得 $x^d - 1 \mid x^r - 1$,此与 $r < d$ 矛盾.故 $r = 0$,即 $d \mid n$.

定理 2.1.2(带余除法)　设 $f(x), g(x)$ 为 $F[x]$ 的任意两个多项式,且 $g(x) \neq 0$,则在 $F[x]$ 中存在唯一的一对多项式 $q(x)$ 和 $r(x)$,使 $f(x) = g(x)q(x) + r(x)$,其中 $r(x) = 0$ 或 $\partial^0 r(x) < \partial^0 g(x)$.称多项式 $q(x)$ 和 $r(x)$ 分别为以 $g(x)$ 除 $f(x)$ 所得的商式和余式.

因此得到两个结论:

1) $g(x) \mid f(x) \Leftrightarrow r(x) = 0$,其中 $r(x)$ 是 $g(x)$ 除 $f(x)$ 所得的余式.

2) 多项式的整除性不因数域扩大而改变.

定义 2.1.4　设 $f(x), g(x), d(x) \in F[x]$,若

1) $d(x) \mid f(x), d(x) \mid g(x)$,即 $d(x)$ 为 $f(x)$ 与 $g(x)$ 的一个公因式;

2) 对 $f(x)$ 与 $g(x)$ 的任一公因式 $c(x)$,都有 $c(x) \mid d(x)$,则称 $d(x)$ 为 $f(x)$ 与 $g(x)$ 的一个**最大公因式**;通常把首项系数为 1 的最大公因式,记为:$(f(x), g(x)) = d(x)$.

定理 2.1.3(最大公因式存在唯一性定理)　$F[x]$ 的任意两个多项式 $f(x)$ 与 $g(x)$ 一定有最大公因式,除相差一个零次因式外,$f(x)$ 与 $g(x)$ 的最大公因式是唯一确定的.

值得注意的是,两个多项式的最大公因式不因数域扩大而改变,但它们的公因式却不然.

定理 2.1.4(倍式和定理)　设 $d(x)$ 是 $F[x]$ 的多项式 $f(x)$ 与 $g(x)$ 的最大公因式,则在 $F[x]$ 中存在多项式 $u(x)$ 与 $v(x)$,使 $f(x)u(x) + g(x)v(x) = d(x)$.

定义 2.1.5　设 $f(x), g(x) \in F[x]$,若 $f(x)$ 与 $g(x)$ 除零次多项式外不再有其他的公因式,则称 $f(x)$ 与 $g(x)$ **互素**,记为:$(f(x), g(x)) = 1$.

上述两个定义可推广到 n 个多项式的情形.值得注意的是,$n(n \geqslant 2)$ 个多项式 $f_1(x), f_2(x) \cdots f_n(x)$ 互素时,它们并不一定两两互素.

定理 2.1.5　$F[x]$ 的两个多项式 $f(x)$ 与 $g(x)$ 互素,即 $(f(x), g(x)) = 1 \Leftrightarrow$ 存在 $u(x), v(x) \in F[x]$,使 $f(x)u(x) + g(x)v(x) = 1$.

互素的性质:

1) 若 $(f(x), h(x)) = 1, (g(x), h(x)) = 1$,则 $(f(x) \cdot g(x), h(x)) = 1$;

2) 若 $h(x) \mid f(x)g(x)$,且 $(f(x), h(x)) = 1$,则 $h(x) \mid g(x)$;

3) 若 $g(x) \mid f(x), h(x) \mid f(x)$,且 $(g(x), h(x)) = 1$,则 $g(x)h(x) \mid f(x)$;

此性质可推广到有限多个多项式的情形.

例 2.1.3　设 $f(x), g(x) \in F[x]$,且 $\partial^0 f(x) > 0, \partial^0 g(x) > 0$,若 $(f(x),$

$g(x))=1$,则在 $F[x]$ 中存在唯一多项式 $u(x),v(x)$,使得

$$f(x)u(x)+g(x)v(x)=1,\text{且 } \partial^0 v(x)<\partial^0 f(x),\partial^0 u(x)<\partial^0 g(x).$$

证明　由于 $(f(x),g(x))=1$,则存在 $h(x),k(x)\in F[x]$,使得 $f(x)h(x)+g(x)k(x)=1$,由带余除法得,$h(x)=g(x)q(x)+u(x),k(x)=f(x)p(x)+v(x)$,其中 $\partial^0 u(x)<\partial^0 g(x)$, $\partial^0 v(x)<\partial^0 f(x)$,由此得,

$$f(x)(g(x)q(x)+u(x))+g(x)(f(x)p(x)+v(x))=1,$$

即

$$f(x)u(x)+g(x)v(x)+f(x)g(x)(q(x)+p(x))=1,$$

如果 $p(x)+q(x)\neq 0$,那么 $\partial^0 f(x)g(x)>\partial^0 f(x)u(x),\partial^0 f(x)g(x)>\partial^0 g(x)v(x)$,等式左端多项式的次数不能为 0,矛盾,所以 $p(x)+q(x)=0$. 故有 $f(x)u(x)+g(x)v(x)=1$.

再证唯一性,假设还有一对满足要求的多项式 $u_0(x),v_0(x)$,使得 $f(x)u_0(x)+g(x)v_0(x)=1$,则有 $(u(x)-u_0(x))f(x)=(v_0(x)-v(x))g(x)$,假设 $v(x)\neq v_0(x)$,则由 $(f(x),g(x))=1$ 推出 $f(x)\mid v(x)-v_0(x)$,但 $\partial^0(v(x)-v_0(x))<\partial^0 f(x)$,矛盾. 故必有 $v(x)=v_0(x),u(x)=u_0(x)$.

2.1.3　不可约多项式

定义 2.1.6　$F[x]$ 中次数大于零且在 F 上只有平凡因式的多项式称为 F 上的**不可约多项式**.

如果次数大于零且在 $F[x]$ 中,有除平凡因式外其他因式的多项式称为在 F 上**可约**.

对零和零次多项式不定义可约性,多项式的可约性与数域有关.

定理 2.1.6　$f(x)$ 在 $F[x]$ 中任一分解式 $f(x)=g(x)h(x)$ 中的因式 $g(x)$ 与 $h(x)$ 总有一个是零次的 $\Leftrightarrow f(x)$ 在 F 上不可约.

不可约多项式的性质:

1) 若 $p(x)$ 为 F 上不可约多项式,则 $cp(x)$ 也是不可约多项式(其中 $0\neq c\in F$);

2) 若 $p(x)$ 为 F 上不可约多项式,$\forall f(x)\in F[x]$,则或 $(f(x),p(x))=1$,或 $p(x)\mid f(x)$;

3) 若 $p(x)$ 为 F 上不可约多项式,且 $p(x)\mid f(x)g(x)$,则 $p(x)\mid f(x)$ 或 $p(x)\mid g(x)$.

例 2.1.4　设 $g(x)=x^5+9x^3+6$,$f(x)$ 在 $\mathbf{Q}$ 上不可约,并且 $f(x)$ 与 $g(x)$ 在复数域内有一公共复根,求证 $f(x)$ 与 $g(x)$ 在 $\mathbf{Q}[x]$ 中相伴,即 $f(x)=cg(x)$,其中 $c(c\neq 0)\in\mathbf{Q}$.

证明　由 Eisenstein 判别法知 $g(x)=x^5+9x^3+6$ 是 $\mathbf{Q}$ 上不可约多项式,而 $f(x)$

也是 $\mathbf{Q}$ 上不可约的，且 $f(x)$ 与 $g(x)$ 在 $\mathbf{C}$ 内有一公共复根，故 $(f(x),g(x)) = d(x) \neq 1$，因此由不可约多项式性质知 $f(x) \mid g(x)$，$g(x) \mid f(x)$，所以 $f(x) = cg(x)$，其中 $c(c \neq 0) \in \mathbf{Q}$.

定理 2.1.7(多项式唯一因式分解定理) $F[x]$ 的每一个 $n(>0)$ 次多项式 $f(x)$ 都可以分解成 $F[x]$ 的不可约多项式的乘积，若不计零次因式的差别，多项式 $f(x)$ 分解成不可约因式乘积的分解式是唯一的. 一般地有：

$$f(x) = ap_1^{k_1}(x)p_2^{k_2}(x)\cdots p_r^{k_r}(x),$$

其中 $p_i(x)(i = 1,2 \cdots r)$ 为 $f(x)$ 的所有互不相同的首项系数为 1 的不可约因式. 上式称为 $f(x)$ 的**典型分解式**.

定义 2.1.7 设 $p(x)$ 是 $F[x]$ 上不可约多项式，且 $p^k(x) \mid f(x)$，但 $p^{k+1}(x)$ 不整除 $f(x)$，则称 $p(x)$ 是 $f(x)$ 的 **k 重因式**. 特别地，当 $k = 1$ 时，称 $p(x)$ 是 $f(x)$ 的**单因式**.

定义 2.1.8 设 $f(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \in F[x]$，$f(x)$ 的导数指 $F[x]$ 的多项式

$$f'(x) = a_1 + 2a_2x + 3a_3x^2 + \cdots + na_nx^{n-1}.$$

上述诸定义都是把多项式看作形式表达式给出的，并且定义 2.1.2～2.1.7 又都限制在数域 F 上的一元多项式环 $F[x]$ 中讨论.

定理 2.1.8 $f(x)$ 无重因式 $\Leftrightarrow (f(x),f'(x)) = 1$.

同时，当 $f(x)$ 的典型分解式形如上式，设 $(f(x),f'(x)) = d(x)$，则

$$q(x) = f(x)/d(x) = a\, p_1(x)\, p_2(x) \cdots p_r(x),$$

即 $q(x)$ 与 $f(x)$ 有完全相同的不可约多项式，且都是单因式.

定理 2.1.9(重因式定理) 设 $p(x)$ 是 $f(x)$ 的一个 $k(k \geqslant 1)$ 重因式，则 $p(x)$ 是 $f'(x)$ 的 $k-1$ 重因式，因而 $p(x)$ 是 $f(x),f'(x) \cdots f^{(k-1)}(x)$ 的公因式，不是 $f^{(k)}(x)$ 的因式.

特别地，当 $k = 1$ 时 $p(x)$ 不是 $f'(x)$ 的因式；反之，若 $p(x) \mid f(x)$，且 $p(x)$ 为 $f'(x)$ 的 $k-1$ 重因式，则 $p(x)$ 为 $f(x)$ 的 k 重因式.

例 2.1.5 数域 F 上的一个 n 次多项式 $f(x)$ 能被它的导数整除 $\Leftrightarrow f(x) = a(x-b)^n$.

证明 先证充分性，因为 $f'(x) = na(x-b)^{n-1}$，所以 $f'(x) \mid f(x)$.

下面给出必要性的证明：

由题设 $f'(x) \mid f(x)$，则有 $g(x) \in F[x]$，使得 $f(x) = g(x)f'(x)$，且 $\partial^0 g(x) = 1$，于是令 $g(x) = a_1(x-b)$，由 $f'(x) \mid f(x)$ 知，$f'(x)$ 是 $f(x)$ 与 $f'(x)$ 的最大公因式. 因此，$g(x) = f(x)/cf'(x)$，与 $f(x)$ 有完全相同的不可约因式，$g(x)$ 只有一个一次因式，故 $f(x) = a(x-b)^n$.

2.1.4　多项式函数

下面我们将从函数观点来讨论多项式.多项式的重要性在于它是最基本的函数,用它可去逼近一个比较复杂的函数,这对数学分析、微分方程等学科,在理论和实际求解上是有重要意义的.

定义 2.1.9　设数环 $\mathbf{R}$ 上一元多项式环 $\mathbf{R}[x]$ 的一个多项式

$$f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n,$$

$c\in\mathbf{R}$,则在 $f(x)$ 的表示式里,把 x 用 c 代替得到 $\mathbf{R}$ 的一个数 $a_0+a_1c+a_2c^2+\cdots+a_nc^n$,称为当 $x=c$ 时 $f(x)$ 的值,用符号 $f(c)$ 表示,于是对于 $\forall c\in\mathbf{R}$,在 $\mathbf{R}$ 中就有唯一确定的数 $f(c)$ 与它对应,这样就得到:$\mathbf{R}$ 到 $\mathbf{R}$ 的一个映射,这个映射是由多项式 $f(x)$ 所确定的,通常称为 $\mathbf{R}$ 上的一个多项式函数.

在此观点下,以及 $\mathbf{R}[x]$ 中 x 又可理解为自变数,对 $\forall c\in\mathbf{R}$,便决定一个 $\mathbf{R}[x]$ 到 $\mathbf{R}$ 的映射 φ: $f(x)\mapsto f(c)$ 是同态映射.

因此,对 $\forall f(x),g(x)\in\mathbf{R}[x]$, $\forall c\in\mathbf{R}$,有

$$f(x)+g(x)\mapsto f(c)+g(c);$$

$$f(x)g(x)\mapsto f(c)g(c).$$

特别地,使 $f(c)=0$ 的 c 称为 $f(x)$ 在 $\mathbf{R}$ 中的根或称 $f(x)$ 的零点.

例 2.1.6　设 $\alpha\in\mathbf{C}$,求证 α 为 $F[x]$ 中某一次数大于零的多项式的根的充要条件是集合 $I=\{f(\alpha)\mid f(x)\in F[x]\}$ 为一数域.

证明　*必要性*　令 α 为 $F[x]$ 中次数大于零的多项式 $f(x)$ 的根,即 $f(\alpha)=0$,则 $0\in I$,因此 $I\neq\varnothing$,同时令 $g(x)=f(x)+1\in F[x]$,则 $1=f(\alpha)+1=g(\alpha)\in I$,并且 $\forall f_1(\alpha)$、$f_2(\alpha)\in I$,由 $f_1(x)$、$f_2(x)\in F[x]$ 推得 $f_1(\alpha)\pm f_2(\alpha)$, $f_1(\alpha)f_2(\alpha)\in I$,故 I 为一数环.

下证数环 I 中每一非零数都有其逆,令

$$J=\{f(x)\in F[x]\mid f(\alpha)=0\},$$

由于 $f(\alpha)=0$,知 $J\neq\varnothing$, J 中必有一不可约多项式 $p(x)$,使 $p(\alpha)=0$,并且 $\forall h(x)\in F[x]$,若 $h(x)\notin J$,即 $h(\alpha)\neq 0$,则 $p(x)$ 不整除 $h(x)$,因此有 $(p(x),h(x))=1$,于是存在 $u(x)$、$v(x)\in F[x]$,使 $p(x)u(x)+h(x)v(x)=1$,当 $x=\alpha$ 时,得 $h(\alpha)v(\alpha)=1$,即 $v(\alpha)=h^{-1}(\alpha)\in I$,故数环 I 为一数域.

充分性　当 $\alpha=0$ 时,则 $f(\alpha)=x$ 即为所求.若 $\alpha\neq 0$,则 α 为 $f(x)=x$ 在 $x=\alpha$ 的值.因此 $\alpha\in I$,由 I 为数域知 $\alpha^{-1}\in I$,于是在 $F[x]$ 中存在一非零多项式 $g(x)$,使 $g(\alpha)=\alpha^{-1}$,即 $\alpha g(\alpha)-1=0$,由此可知 α 为多项式 $xg(x)-1$ 的根.

又因为 $g(x) \neq 0$,所以 $xg(x)-1$ 的次数大于零,令 $f(x)=xg(x)-1$ 即为所求.

定理 2.1.10(余式定理) 设 $f(x) \in \mathbf{R}[x], c \in \mathbf{R}$,用 $x-c$ 除 $f(x)$ 所得的余式 $r=f(c)$.

定理 2.1.11(因式定理) $x-c \mid f(x) \Leftrightarrow f(c)=0$.

2.1.5 复、实、有理数域上的不可约多项式

实系数多项式的非实复根两两成对,因此实数域上不可约多项式,除一次多项式外,只有含非实共轭复根的二次多项式;每一实系数多项式都可分解为实系数的一次和二次不可约因式的乘积.

根与系数的关系:设 $\alpha_1, \alpha_2 \cdots \alpha_n$ 为 n 次多项式 $f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n$ 的 n 个根,则

$$\alpha_1+\alpha_2+\cdots+\alpha_n=-a_1/a_0;$$

$$\alpha_1\alpha_2+\alpha_1\alpha_3+\cdots+\alpha_{n-1}\alpha_n=a_2/a_0;$$

$$\alpha_1\alpha_2\alpha_3+\alpha_1\alpha_2\alpha_4+\cdots+\alpha_{n-2}\alpha_{n-1}\alpha_n=(-1)^3a_3/a_0;$$

$$\cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots \quad \cdots$$

$$\alpha_1\alpha_2 \quad \cdots \quad \alpha_n=(-1)^n a_n/a_0.$$

复数域 $\mathbf{C}$ 上只有 1 次多项式是不可约的.

定理 2.1.12(代数基本定理) 任何 $n(n>0)$ 次多项式在复数域中至少有一个根.因此任何 $n(n>0)$ 次多项式在复数域中有 n 个根(重根按重数计算).

定义 2.1.10 令 $f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n \in F[x]$ 满足 $(a_0,a_1,\cdots,a_n)=1$(即它的系数互素)称为**本原多项式**.

定理 2.1.13(Gauss 引理) 两个本原多项式的乘积仍是一个本原多项式.

定理 2.1.14 整系数多项式在 $\mathbf{Z}$ 上可约 $\Leftrightarrow$ 它在 $\mathbf{Q}$ 上可约.

有理数域上存在任意次数的不可约多项式.

定理 2.1.15(Eisenstein 判别法) 设 $f(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}+a_nx^n$ 为一整系数多项式;若能找到一个素数 p,使

1) p 不整除 a_n;

2) $p \mid a_i(i=0,1\cdots n-1)$;

3) p^2 不整除 a_0;

则 $f(x)$ 在有理数域 $\mathbf{Q}$ 上不可约.

定理 2.1.16 设有理数 $\frac{u}{v}$ 是整系数多项式 $f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n$ 的一个根,其中 $(u,v)=1$,且 $a_0a_n \neq 0$,则

1) $v \mid a_0, u \mid a_n$;

2) $f(x)=\left(x-\frac{u}{v}\right)q(x)$，其中 $q(x)$ 为整系数多项式.

例 2.1.7　设 $a_1,a_2\cdots a_n$ 是互不相同的整数，则 $f(x)=\prod_{i=1}^{n}(x-a_i)-1$ 在 **Q** 上可约.

证明　反证法　设 $f(x)$ 可约，则 $f(x)$ 可分解为两个整系数多项式的积，即 $f(x)=g(x)h(x)$，又 $f(a_i)=-1$，即 $g(a_i)h(a_i)=-1$，则 $g(a_i)$ 与 $h(a_i)$ 中一个等于 -1，一个等于 0，所以 $g(a_i)+h(a_i)=0$，因 $\partial^0 gf(x)<n,\partial^0 h(x)<n,a_1,a_2\cdots a_n$ 是互不相同的整数，所以 $g(x)+h(x)=0$，故 $g(x)=-h(x)$，可得 $f(x)=-g^2(x)$，这与 $f(x)$ 为首一多项式矛盾.

例 2.1.8　设 $A\in M_n(\mathbf{C}),f(x)\in\mathbf{C}[x]$，且 $\partial^0 f(x)>0,g(x)$ 是以 A 为根的次数最低的多项式，求证：

1) 若 $(f(x),g(x))=d(x)$，则 $d(A)$ 的秩与 $g(A)$ 的秩相等；

2) $f(A)$ 可逆 $\Leftrightarrow(f(x),g(x))=1$.

证明　1) 设 $d(x)=(f(x),g(x))$，则有 $f(x)u(x)+g(x)v(x)=d(x)$，于是有

$$f(A)u(A)+g(A)v(A)=d(A),$$

由 $g(A)=0$，得 $f(A)u(A)=d(A)$，因此有，秩 $f(A)\geqslant$ 秩 $d(A)$；

另一方面，由 $d(x)\mid f(x)$，则有 $f(x)=d(x)q(x)$，于是 $f(A)=d(A)q(A)$，故又有秩 $d(A)\geqslant$ 秩 $f(A)$. 所以 秩 $d(A)=$ 秩 $f(A)$.

2) 必要性　由题设知，$(f(x),g(x))=d(x)$，假若 $\partial^0 d(x)\neq 0$，则由 $g(x)=d(x)g_1(x)$，有 $d(A)g_1(A)=g(A)=0$. 因 $f(A)$ 可逆，由(1) 知 $d(A)$ 也可逆，故 $g_1(A)=0$，此与 $g(x)$ 的题设矛盾，所以 $d(x)$ 必为非零常数.

充分性　若 $(f(x),g(x))=1$，则 $f(x)u(x)+g(x)v(x)=1$，进而有 $f(A)u(A)+g(A)v(A)=I$，由 $g(A)=0$ 得，$f(A)u(A)=I$，故 $f(A)$ 可逆.

2.2　多元多项式

2.2.1　多元多项式的定义及运算

在这一节里我们先介绍多元多项式的概念和多元多项式的一些最基本的性质，然后讨论一下在理论和应用上都比较重要的一种多元多项式——对称多项式.

定义 2.2.1　令 $x_1,x_2,\cdots,x_n$ 是 n 个文字. 形如 $ax_1^{k_1}x_2^{k_2}\cdots x_n^{k_n}$ 的表示式(其中 $a\in\mathbf{R}$，$k_1,k_2,\cdots,k_n$ 是非负整数) 叫做 **R** 上 $x_1,x_2,\cdots,x_n$ 的一个**单项式**.

数 a 叫做这个单项式的系数. 如果某一 $k_i=0$，那么 $x_i^{k_i}$ 可以不写，因此，$m(m<n)$

个文字的单项式总可以看成 n 个文字的单项式. 特别, $0 \in \mathbf{R}$.

一些(有限个)单项式用加号联结起来而得到的一个形式表达式

$$a_1 x_1^{k_{11}} x_2^{k_{12}} \cdots x_n^{k_{1n}} + a_2 x_1^{k_{21}} x_2^{k_{22}} \cdots x_n^{k_{2n}} + \cdots + a_s x_1^{k_{s1}} x_2^{k_{s2}} \cdots x_n^{k_{sn}}, \tag{2.2}$$

其中 $a_i \in \mathbf{R}, k_{ij}$ 是非负整数($i = 1, \cdots, s; j = 1, \cdots, n$),叫做 $\mathbf{R}$ 上 n 个文字 $x_1, x_2, \cdots, x_n$ 的一个多项式,或简称 $\mathbf{R}$ 上一个 n 元多项式,也可以简称为多项式.

定义 2.2.2 $\mathbf{R}$ 上两个 n 元**多项式相等**指的是,如果它们有完全相同的项,或者只差一些系数为零的项.

定义 2.2.3 设 $ax_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ 是 $\mathbf{R}$ 上 n 个文字的一个单项式,且 $a \neq 0$,则非负整数 $k_1 + k_2 + \cdots + k_n$ 叫做这个单项式的次数.

$\mathbf{R}$ 上一个 n 元多项式的次数指的是系数不为零的项的次数最大者.

根据这个定义,零多项式,也就是系数都是零的多项式,是唯一没有定义次数的多项式,如果 $0 \neq a \in \mathbf{R}$,那么 a 是一个零次多项式.

定义 2.2.4 $\mathbf{R}$ 上两个 n 元多项式 $f(x_1, x_2, \cdots, x_n)$ 与 $g(x_1, x_2, \cdots, x_n)$ 的和指的是把分别出现在这两个多项式中对应的同类项的系数相加所得到的 n 元多项式,记作 $f(x_1, x_2, \cdots, x_n) + g(x_1, x_2, \cdots, x_n)$,或者简单地记作 $f + g$.

定义 2.2.5(两个单项式的乘积) $\mathbf{R}$ 上两个 n 元单项式 $ax_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$ 与 $bx_1^{h_1} x_2^{h_2} \cdots x_n^{h_n}$ 的积指的是单项式 $abx_1^{k_1+h_1} x_2^{k_2+h_2} \cdots x_n^{k_n+h_n}$.

设 f 与 g 都是 $\mathbf{R}$ 上 n 个文字 $x_1, x_2, \cdots, x_n$ 的多项式. 把 f 的每一项与 g 的每一项相乘,然后把这些乘积相加(合并同类项)而得到的一个 n 元多项式叫做 f 与 g 的积,记作 fg.

2.2.2 数环 R 上 n 元多项式环

定义 2.2.6 我们把一个数环 $\mathbf{R}$ 上一切 n 个文字 $x_1, x_2, \cdots, x_n$ 的多项式所成的集合,连同如上定义的加法和乘法叫做 $\mathbf{R}$ 上 n 个文字 $x_1, x_2, \cdots, x_n$ 的多项式环,简称 **$\mathbf{R}$ 上 n 元多项式环**,记作 $\mathbf{R}[x_1, x_2, \cdots, x_n]$.

由以上的讨论可知,如果 $\{i_1, i_2, \cdots, i_n\}$ 是 $\{1, 2, \cdots, n\}$ 的一个子集,那么

$$\mathbf{R}[x_{i_1}, x_{i_2} \cdots x_{i_n}] \subseteq \mathbf{R}[x_1, x_2, \cdots, x_n].$$

特别地,对于任意 $n > 0$, $\mathbf{R} \subseteq \mathbf{R}[x_1, x_2, \cdots, x_n]$.

设 $f \in \mathbf{R}[x_1, x_2, \cdots, x_n]$,把 f 的系数都换成各自的相反数,可得到 $\mathbf{R}[x_1, x_2, \cdots, x_n]$ 的一个多项式,叫做 f 的负多项式,记作 $-f$. 显然有 $f + (-f) = 0$.

当 $n = 2$ 时,数环 $\mathbf{R}$ 上二元多项式的一个表达式: $f = \sum a_{ij} x^i y^j$,这里 $a_{ij} \in \mathbf{R}$,其中只有有限个 a_{ij} 不为零, i, j 都是非负整数($\sum$ 表示有限项的和). 我们把 $i + j$ 等于一

个固定整数的项放在一起，那么 f 可以写成以下形式：

$$f = a_{00} + (a_{10}x + a_{01}y) + (a_{20}x^2 + a_{11}xy + a_{02}y^2) + (a_{30}x^3 + a_{21}x^2y + a_{12}xy^2 + a_{03}y^3) + \cdots,$$

在这里自然只有有限项不为零.

采用这种记号，数环 $\mathbf{R}$ 上 n 元多项式的一般形式为 $\sum a_{i_1 i_2 \cdots i_n} x_1^{i_1} x_2^{i_2} \cdots x_n^{i_n}$.

现在来讨论两个 n 元多项式的和与积的次数问题.

设 f,g 是 $\mathbf{R}$ 上两个不等于零的 n 元多项式，如同一元多项式的情形一样，我们有 $\partial^0(f+g) \leqslant \max(\partial^0 f, \partial^0 g)$. 设 $f(x_1, x_2, \cdots, x_n)$ 是数环 $\mathbf{R}$ 上一个不等于零的 n 元多项式. 设

$$(1)\ ax_1^{k_1}x_2^{k_2}\cdots x_n^{k_n} \quad (a \neq 0), \qquad (2)\ bx_1^{l_1}x_2^{l_2}\cdots x_n^{l_n} \quad (b \neq 0).$$

是 $f(x_1, x_2, \cdots, x_n)$ 的两个不同的项，那么在这两项对应的幂指数的差 $k_i - l_i (1 \leqslant i \leqslant n)$ 中，至少有一个不等于零. 如果在这些差中，第一个不等于零的数是一个正数，换句话说，如果存在这样一个 $i, 1 \leqslant i \leqslant n$，使得 $k_1 = l_1, k_{i-1} = l_{i-1}$，但 $k_i > l_i$，那么就说，项(1)大于项(2)，或者项(2)小于项(1). 对于 $f(x_1, x_2, \cdots, x_n)$ 的任意两个不同的项，总有一个大于另一个，并且若项(1)大于项(2)，而项(2)又大于另外一项：(3) $cx_1^{m_1}x_2^{m_2}\cdots x_n^{m_n} (c \neq 0)$，那么项(1)大于项(3). 这样，只要把两项中较大的一项排在前面，多项式 $f(x_1, x_2, \cdots, x_n)$ 的各项就有了完全确定的次序. 这种排列多项式的项的方法很像字典里字的排列法，所以通常把这种排列法叫做多项式的字典排列法.

例如，$f(x_1, x_2, x_3, x_4) = x_1^4 + 3x_1^2x_2^3x_3 - x_1^2x_2^3x_4^2 + x_3^2x_4 - 2$.

这就是按字典排列法书写的一个四元多项式.

定理 2.2.1　数环 $\mathbf{R}$ 上两个 n 元多项式 $f(x_1, x_2, \cdots, x_n)$ 与 $g(x_1, x_2, \cdots, x_n)$ 的乘积的首项等于这两个多项式首项的乘积.

推论 2.2.2　如果 $f_i \neq 0, i = 1, 2, \cdots, n$，则 $f_1 f_2 \cdots f_n$ 的次数等于每个 f_i 的首项的乘积.

推论 2.2.3　两个非零多项式的乘积也不等于零.

定理 2.2.4　数环 $\mathbf{R}$ 上两个非零 n 元多项式的乘积的次数等于它们次数的和.

证明　设 $f,g \in \mathbf{R}[x_1, \cdots, x_n]$，且 $f \neq 0, g \neq 0$，它们的次数分别是 s 和 t，把 f 与 g 写成齐次多项式的和：$f = f_0 + f_1 + \cdots + f_s, g = g_0 + g_1 + \cdots + g_t$，这里 f_i, g_i 或者等于零，或者分别是 i 次或 j 次齐式($i = 0, \cdots, s; j = 0, \cdots, t$)，并且 $f_s \neq 0, g_t \neq 0$，于是，

$$fg = f_0 g_0 + f_0 g_1 + f_1 g_0 + \cdots + f_s g_t + \cdots$$

由定理 2.2.1，$f_s g_t \neq 0$，并且是一个 $s+t$ 次齐式，而其余各项 $f_i g_i$ 或者等于零，或者是一个次数低于 $s+t$ 的齐式. 因此，

$$\partial^0(fg) = s + t = \partial^0 f + \partial^0 g.$$

定义 2.2.7 数环 **R** 上一个 n 元多项式 $f(x_1, \cdots, x_n)$ 和 **R** 中任意 n 个数 $c_1, \cdots, c_n$ 在 $f(x_1, \cdots, x_n)$ 中,把 x_i 用 c_i 来代替,就得到数环 **R** 的一个确定的数,叫做 $x_i = c_i(i=1,\cdots, n)$ 时多项式 $f(x_1,\cdots, x_n)$ 的**值**,并且用符号 $f(c_1,\cdots, c_n)$ 来表示. 如果 $f(c_1, \cdots, c_n) = 0$,那么数组$(c_1,\cdots, c_n)$叫做 $f(x_1, \cdots, x_n)$ 的一个**零点**.

2.2.3 *n* 元多项式与多项式函数

定义 2.2.8 设 **R** 是一个数环. 因为假定了 $1 \in \mathbf{R}$,所以 **R** 含有无限多个数. 令 $R^n = \{(c_1, c_2, \cdots, c_n) \mid c_i \notin \mathbf{R}, 1 \leqslant i \leqslant n\}$. 给定了 R 上一个 n 元多项式 $f(x_1,\cdots, x_n)$.

可以定义一个函数(映射)$R^n \to R$,使得$(c_1,\cdots, c_n) \mapsto f(c_1, \cdots, c_n)$,这个函数称为由多项式 $f(x_1, x_2, \cdots, x_n)$ 所确定的多项式函数.

设 $f(x_1, \cdots, x_n), g(x_1, \cdots, x_n) \in \mathbf{R}[x_1, \cdots, x_n]$.

定义 2.2.9 在 $\mathbf{R}[x_1,\cdots, x_n]$ 中,如果 $f(x_1, \cdots, x_n) = g(x_1, \cdots, x_n)$,那么对于 $\forall(c_1,\cdots,c_n) \in R^n$ 都有 $f(c_1, \cdots, c_n) = g(c_1,\cdots, c_n)$, 即由 $f(x_1,\cdots, x_n)$ 和 $g(x_1,\cdots, x_n)$ 所确定的多项式函数 f 与 g 相等.

我们要证明这一事实的反面也成立. 为此,只需证明

定理 2.2.5 设 $f(x_1,\cdots, x_n)$ 是数环 **R** 上一个 n 元多项式. 如果对于任意$(c_1,\cdots, c_n) \in R^n$ 都有 $f(c_1,\cdots, c_n) = 0$, 那么 $f(x_1,\cdots, x_n) = 0$.

证明 显然当 $n = 1$ 时结论成立.

假设 $n > 1$,并且对于 **R** 上 $n-1$ 个文字的多项式来说定理成立. 设 $f(x_1,\cdots, x_n)$ 是 **R** 上一个 n 元多项式. 把含有 x_n 同一次幂的项放在一起,把 x_n 的幂提到括号外边,那么 $f(x_1,\cdots, x_n)$ 可以写成

$$f(x_1, \cdots, x_n) = u_0 + u_1 x_n + \cdots + u_t x_n^t,$$

这里 $u_i = (x_1,\cdots, x_{n-1}) \in \mathbf{R}(x_1, \cdots, x_n)$, $i = 0,1, \cdots, t$. 任取 $c = (c_1, \cdots,c_{n-1}) \in R^{n-1}$. 在每一 $u_i(i = 0, \cdots, t)$ 里,以 c_j 代替 $x_j(j = 1, \cdots, n-1)$,我们得到 **R** 上一个文字 x_n 的多项式 $g_c(x_n) = u_0(c) + u_1(c)x_n + \cdots + u_t(c)x_n^t$,这里 $u_i(c) = u_i(c_1,c_2, \cdots,c_{n-1}) \in \mathbf{R}$,如果对于任意$(c_1, \cdots,c_{n-1}) \in \mathbf{R}$ 都有 $f(c_1, \cdots,c_n) = 0$,那么 $c = (c_1, \cdots,c_{n-1}) \in R^{n-1}$ 和 $c \in \mathbf{R}$,我们有

$$g_c(x_n) = f(c_1,c_2, \cdots,c_n) = u_0(c) + u_1(c)c_n + \cdots + u_t(c)c_n^t = 0,$$

因定理对于一元多项式成立,所以 $g_c(x_n) = 0$,即

$$u_i(c) = u_i(c_1,c_2, \cdots,c_{n-1}), i = 1,2, \cdots, t,$$

然而 $c = (c_1, \cdots, c_{n-1})$ 可以取遍 R^{n-1} 中一切元素. 由归纳法假定,必须

$$u_i = (x_1, x_2, \cdots, x_{n-1}) = 0, i = 1, \cdots, t,$$

从而 $f(x_1, \cdots, x_n) = 0$.

推论 2.2.6　设 $f(x_1, \cdots, x_n)$、$g(x_1, \cdots, x_n) \in \mathbf{R}[x_1, x_2, \cdots, x_n]$. 如果对于 $\forall (c_1, \cdots, c_n) \in R^n$,有 $f(c_1, \cdots, c_n) = g(c_1, \cdots, c_n)$,则 $f(x_1, \cdots, x_n) = g(x_1, \cdots, x_n)$.

推论 2.2.7　如果由 $f(x_1, x_2, \cdots, x_n)$ 与 $g(x_1, x_2, \cdots, x_n)$ 确定的多项式函数 f 与 g 相等,那么这两个多项式相等.

2.2.4　对称多项式

在这一节里,我们将讨论对称多项式. 对称多项式是一类重要的多元多项式,它的应用也比较广泛.

定义 2.2.10　设 $f(x_1, x_2, \cdots, x_n)$ 是数环 $\mathbf{R}$ 上一个 n 元多项式. 如果对于这 n 个文字 $x_1, x_2, \cdots, x_n$ 的指标集$\{1,2, \cdots, n\}$ 施行任意一个置换后, $f(x_1, x_2, \cdots, x_n)$ 都不改变,那么就称 $f(x_1, x_2, \cdots, x_n)$ 是 $\mathbf{R}$ 上一个 n 元对称多项式.

例 2.2.1　$f(x_1, x_2, \cdots, x_n) = x_1^2 + x_2^2 + \cdots + x_n^2$ 是整数环上一个 n 元对称多项式. $f(x_1, x_2, x_3) = x_1^2x_2 + x_1^2x_3 + x_1x_2^2 + x_1x_3^2 + x_2^2x_3 + x_2x_3^2$ 是 $\mathbf{Z}$ 上一个三元对称多项式.

不难看出,两个 n 元对称多项式的和、差、积仍是 n 元对称多项式.

在对称多项式理论中,所谓初等对称多项式占很重要的地位.

定义 2.2.11　有以下 n 个 n 元对称多项式:

$$\sigma_1 = x_1 + x_2 + \cdots + x_n$$

$$\sigma_2 = x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n$$

$$\cdots\cdots\cdots\cdots$$

$$\sigma_{n-1} = x_1x_2\cdots x_{n-1} + x_1x_2\cdots x_{n-2}x_n + \cdots + x_2x_3\cdots x_n$$

$$\sigma_n = x_1x_2\cdots x_n.$$

称这 n 个多项式 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 为 n 元**初等对称多项式**.

定理 2.2.8　设 $f(x_1, x_2, \cdots, x_n) = \sum a_{i_1i_2\cdots i_n} x_1^{i_1}x_2^{i_2}\cdots x_n^{i_n}$ 是数环 $\mathbf{R}$ 上一个 n 元多项式. 以 σ_i 代替 x_i, $1 \leqslant i \leqslant n$,得到关于 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的一个多项式

$$f(\sigma_1, \sigma_2, \cdots, \sigma_n) = \sum a_{i_1i_2\cdots i_n}\sigma_1^{i_1}\sigma_2^{i_2}\cdots\sigma_n^{i_n}.$$

如果 $f(\sigma_1, \sigma_2, \cdots, \sigma_n) = 0$,那么一切系数 $a_{i_1i_2\cdots i_n} = 0$,即 $f(x_1, x_2, \cdots, x_n) = 0$.

定理 2.2.9　数环 $\mathbf{R}$ 上每一 n 元对称多项式 $f(x_1, x_2, \cdots, x_n)$ 都可以表成初等对称多项式 $\sigma_1, \sigma_2, \cdots, \sigma_n$ 的系数在 $\mathbf{R}$ 中的多项式,并且这种表示法是唯一的.

基本定理的证明给了一个实际用初等对称多项式来表示对称多项式的方法.

例 2.2.2 用初等对称多项式表示 n 元对称多项式 $f = x_1^2 + x_2^2 + \cdots + x_n^2$.

解 f 的首项是 x_1^2,所以取 $g_1 = \sigma_1^{2-0}\sigma_2^0\cdots\sigma_n^0 = \sigma_1^2$,于是

$$f_1 = f - g_1 = x_1^2 + x_2^2 + \cdots + x_n^2 = -(x_1 + x_2 + \cdots + x_n)^2$$

$$= -2(x_1x_2 + x_1x_3 + \cdots + x_{n-1}x_n) = -2\sigma_2.$$

所以 $f = g_1 + f_1 = \sigma_1^2 - 2\sigma_2$.

定理 2.2.10 设 $f(x) = x^n + a_1x^{n-1} + \cdots + a_{n-1}x + a_n$ 是某一数域 F 上一个多项式.令 $\alpha_1,\cdots,\alpha_n$ 是 $f(x)$ 在 $\mathbf{C}$ 内的全部根,则 $-\alpha_1, \alpha_2, \cdots, (-1)^n\alpha_n$ 就是 $\alpha_1, \cdots, \alpha_n$ 的初等对称多项式.

定理 2.2.11 设 $f(x)$ 是数域 F 上一个一元 n 次多项式,它的最高次项系数是 1.令 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 是 $f(x)$ 在复数域内的全部根(重根按重数计算),那么 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的每一个系数取自 F 的对称多项式都是 $f(x)$ 的系数的多项式,因而是 F 的一个数.

2.3 行列式的计算

行列式是线性代数的重要组成部分,它最早出现在 16 世纪关于线性方程组的求解问题中,后来行列式理论发展很快,它在消元法、矩阵论、二次型及二次型简化为标准型、多重积分中的变量替换、微分方程组等问题中都有广泛的应用,这些应用最终都离不开行列式的计算.掌握行列式的概念和行列式的性质,并会灵活地计算行列式,提高行列式的计算技巧是本节内容的重点.

2.3.1 n 阶行列式的定义

定义 2.3.1 n^2 个数,排成 n 行,n 列,即

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

等于 $n!$ 项的代数和,其中每一项都是取自不同行不同列的 n 个元素的乘积 $a_{1j_1}a_{2j_2}\cdots a_{nj_n}$,而其符号为 $(-1)^{\pi(j_1j_2\cdots j_n)}$.此定义可记为

$$D_n = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = \sum_{j_1j_2\cdots j_n}^{n!} (-1)^{\pi(j_1j_2\cdots j_n)} a_{1j_1}a_{2j_2}\cdots a_{nj_n}.$$

定义 2.3.2(箭形行列式)　形如

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & \cdots & a_{1n} \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & 0 & a_{33} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ a_{n1} & 0 & \cdots & 0 & a_{nn} \end{vmatrix}$$

简记为 $|\nwarrow|$ 或 $|\nearrow|$，$|\swarrow|$，$|\searrow|$，统称为箭形行列式.

2.3.2　行列式基本性质

(1) 行到式与它的转置行列式相等，即 $D' = D$.

(2) 交换行列式的两行(或两列)，行列式改变符号，即 $D \xrightarrow{(i,j)} D_1 = -D$.
由此即得，若行列式某两行(或列) 完全相同，则此行列式为零.

(3) 把一个行列式的某一行(或列) 的所有元素同乘以某一数 k，等于以数 k 乘这个行列式，即 $D \xrightarrow{k(i)} D_1 = kD$. 由此可得三个结论：

1) 行列式中某一行(列) 所有元素的公因子可提到行列式外边来；

2) 若行列式中有一行(列) 的元素全都是零，则此行列式等于零；

3) 若行列式中有两行(列) 的对应元素成比例，则此行列式等于零.

(4) 设行列式 D 的第 i 行的所有元素都可表为两元素的和 $a_{ij} = b_{ij} + c_{ij}$，$j = 1,\cdots,n$

$$D = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ b_{i1}+c_{i1} & b_{i2}+c_{i2} & \cdots & b_{in}+c_{in} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix},$$

那么 $D = D_1 + D_2$，其中 D_1 与 D_2 的第 i 行的元素分别是 $b_{i1}\cdots b_{in}$，$c_{i1}\cdots c_{in}$，而它们的其余各行与 D 的对应行相同. 对列同样的性质也成立.

(5) 把行列式的某一行(列) 的元素乘以同一数后加到另一行(列) 的对应元素上，行列式不变，即 $D \xrightarrow{k(j)+(i)} D_1 = D$.

2.3.3　代数余子式

定义 2.3.3　在一个 n 阶行列式 D 中任选 k 行与 k 列($1 \leqslant k \leqslant n$)，位于这些行和列相交处的 k^2 个元素按原次序组成一个 k 阶行列式 M 称为行列式 D 的一个 **k 阶**

子式；

在一个 n 阶行列式 D 中划去它的一个 k 阶子式 M 所在的 k 行和 k 列（$1 \leqslant k \leqslant n$）后余下的元素按原次序组成的 $n-k$ 阶行列式 M' 称为 k 阶子式 M 的**余子式**.

特别地，当 $k=1$ 时，即 M 为一阶子式 a_{ij}，把元素 a_{ij} 所在的第 i 行与第 j 列划去余下的 $n-1$ 行与 $n-1$ 列按原次序构成的 $n-1$ 阶子式便是元素 a_{ij} 的余子式 M_{ij}；元素 a_{ij} 的余子式附上符号 $(-1)^{i+j}$ 便为 a_{ij} 的代数余子式 $A_{ij}=(-1)^{i+j}M_{ij}$.

定理 2.3.1（行列式依行依列的展开定理）

$$a_{i1}A_{j1}+a_{i2}A_{j2}+\cdots+a_{in}A_{jn}=\begin{cases}D\cdots i=j\\0\cdots i\neq j\end{cases}, a_{1i}A_{1j}+a_{2i}A_{2j}+\cdots+a_{ni}A_{nj}=\begin{cases}D\cdots i=j\\0\cdots i\neq j\end{cases}$$

2.3.4　几个重要定理

定理 2.3.2（Cramer 规则）　若一个含有 n 个未知量 n 个方程的线性方程组

$$\begin{cases}a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1\\a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2\\\cdots\cdots\cdots\\a_{n1}x_1+a_{n2}x_2+\cdots+a_{nn}x_n=b_n\end{cases}\tag{2.3}$$

的系数行列式 $D=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}\neq 0$，

则线性方程组(2.3)有且仅有一个解：$x_j=\dfrac{D_j}{D}$，$j=1,2\cdots n$，

其中 D_j 是 D 中第 j 列元素换以 $b_1, b_2, \cdots, b_n$ 而得到的 n 阶行列式.

定理 2.3.3［拉普拉斯（Laplace）定理］　设在 n 阶行列式 D 中任意取定 k（$1\leqslant k\leqslant n-1$）个行（或列），那么由这 k 行（或列）元素所组成的一切 k 阶子式与它们的代数余子式的乘积之和等于行列式 D.

定义 2.3.4　两个 n 阶行列式

$$D_1=\begin{vmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{n1}&a_{n2}&\cdots&a_{nn}\end{vmatrix}\text{ 与 }D_2=\begin{vmatrix}b_{11}&b_{12}&\cdots&b_{1n}\\b_{21}&b_{22}&\cdots&b_{2n}\\\vdots&\vdots&&\vdots\\b_{n1}&b_{n2}&\cdots&b_{nn}\end{vmatrix}$$

的乘积 $D_1 \cdot D_2 = \begin{vmatrix} c_{11} & c_{12} & \cdots & c_{1n} \\ c_{21} & c_{22} & \cdots & c_{2n} \\ \vdots & \vdots & & \vdots \\ c_{n1} & c_{n2} & \cdots & c_{nn} \end{vmatrix}$,

其中 $c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{in}b_{nj}, i,j = 1,2 \cdots n$.

这一乘法规则符合两个矩阵的乘法规则.

定理 2.3.4[第一降阶定理(Schur 定理)]

设 $\begin{pmatrix} A & B \\ C & D \end{pmatrix}$ 是方阵,且 A 可逆,则 $\begin{vmatrix} A & B \\ C & D \end{vmatrix} = |A||D - CA^{-1}B|$.

定理 2.3.5(第二降阶定理)　设 A 与 D 分别是 n 阶和 m 阶可逆阵,B,C 分别是 $n \times m$ 和 $m \times n$ 矩阵,则

$$|D - CA^{-1}B| = |D| / |A| \; |A - BD^{-1}C|;$$

$$|D + CA^{-1}B| = |D| / |A| \; |A + BD^{-1}C|.$$

2.3.5　行列式计算

1. 利用行列式的定义

例 2.3.1　求 $\sum\limits_{t_1 t_2 \cdots t_n} \begin{vmatrix} a_{1t_1} & a_{1t_2} & \cdots & a_{1t_n} \\ a_{2t_1} & a_{2t_2} & \cdots & a_{2t_n} \\ \vdots & \vdots & & \vdots \\ a_{nt_1} & a_{nt_2} & \cdots & a_{nt_n} \end{vmatrix}$ 的值.

解　排列 $t_1, t_2 \cdots t_n$ 经过与排列反序数相同次数的对换化为 $1,2 \cdots n$ 的排列,又由于所有 n 元排列中奇、偶排列各一半,所以在原行列式中一半经过奇数次对换,一半经过偶数次对换后化为

$$\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

也就是说,在 $n!$ 个行列式中一半为正,一半为负,所以总和为 0.

2. 利用行列式性质

利用行列式性质,把原行列式化为上(下) 三角形,这样原行列式等于此上(下) 三角形行列式的主对角线元素之积.

例 2.3.2 计算行列式 $D_{n+1}=\begin{vmatrix} a_1 & -a_1 & & & \\ & a_2 & -a_2 & & \\ & & \ddots & \ddots & \\ & & & a_n & -a_n \\ b & b & \cdots & b & b \end{vmatrix}$.

解 各列加到第一列得：

$$D_{n+1}=\begin{vmatrix} 0 & -a_1 & & & \\ 0 & a_2 & -a_2 & & \\ \vdots & & \ddots & \ddots & \\ 0 & & & a_n & -a_n \\ (n+1)b & b & \cdots & b & b \end{vmatrix}$$

$$=(-1)^{(n+1)+1}(n+1)b(-a_1)(-a_2)\cdots(-a_n)$$

$$=(n+1)a_1a_2\cdots a_nb.$$

3. 化为箭形行列式

例 2.3.3 计算行列式 $D_n=\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 0 & \cdots & 0 \\ 1 & 0 & 3 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & n \end{vmatrix}$.

解 $D_n \xrightarrow{-\frac{1}{i}(i)+(1)} \begin{vmatrix} 1-\sum\limits_{i=2}^{n}\frac{1}{i} & 0 & 0 & \cdots & 0 \\ 1 & 2 & 0 & \cdots & 0 \\ 1 & 0 & 3 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 1 & 0 & \cdots & 0 & n \end{vmatrix} = n!\left(1-\sum\limits_{i=2}^{n}\frac{1}{i}\right).$

例 2.3.4 计算行列式 $D_n=\begin{vmatrix} a_1 & x & \cdots & x \\ x & a_2 & \cdots & x \\ \vdots & \vdots & \ddots & \vdots \\ x & x & \cdots & a_n \end{vmatrix}$.

解 把第一行各元素皆乘以 -1 分别加到各行对应元素上，得到

$$\begin{vmatrix} a_1 & x & x & \cdots & x \\ x-a_1 & a_2-x & 0 & \cdots & 0 \\ x-a_1 & 0 & a_3-x & \ddots & \vdots \\ \vdots & \vdots & \vdots & \ddots & 0 \\ x-a_1 & 0 & \cdots & 0 & a_n-x \end{vmatrix},$$

对此行列式各列分别提取公因式 $a_1-x, a_2-x \cdots a_n-x$，于是得

$$D_n=\prod_{i=1}^{n}(a_i-x)\begin{vmatrix} \frac{a_1}{a_1-x} & \frac{x}{a_2-x} & \cdots & \cdots & \frac{x}{a_n-x} \\ -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -1 & 0 & \cdots & 0 & 1 \end{vmatrix},$$

各列加到第一列上得

$$=\prod_{i=1}^{n}(a_i-x)\begin{vmatrix} \frac{a_1}{a_1-x}+\sum_{j=2}^{n}\frac{x}{a_j-x} & \frac{x}{a_2+x} & \cdots & \cdots & \frac{x}{a_n-x} \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \end{vmatrix}$$

$$=x\prod_{i=1}^{n}(a_i-x)\left(\frac{1}{x}+\sum_{j=1}^{n}\frac{1}{a_j-x}\right).$$

4. *递推法*

设 D_n 的递推关系式 $D_n=pD_{n-1}+qD_{n-1}$，$(n>2)$，其中 p、q 为常数，且 $q\neq 0$，若 α、β 为方程 $x^2-px-q=0$ 的根，则

$$D_n=\frac{1}{\alpha-\beta}[\alpha^{n-1}(D_2-\beta D_1)-\beta^{n-1}(D_2-\alpha D_1)].$$

证明　由于 α、β 为方程 $x^2-px-q=0$ 的根，则 $\alpha+\beta=p$，$\alpha\beta=-q$，因此，

$D_n=(\alpha+\beta)D_{n-1}-\alpha\beta D_{n-2}$；$D_n-\alpha D_{n-1}=\beta(D_{n-1}-\alpha D_{n-2})$；$D_n-\beta D_{n-1}=\alpha(D_{n-1}-\beta D_{n-2})$；

当 $\alpha\neq\beta$ 时，$D_n-\alpha D_{n-1}=\beta(D_{n-1}-\alpha D_{n-2})=\cdots=\beta^{n-2}(D_2-\alpha D_1)$，

同理　$D_n-\beta D_{n-1}=\alpha^{n-2}(D_2-\beta D_1)$，

故　$D_n=\dfrac{1}{\alpha-\beta}[\alpha^{n-1}(D_2-\beta D_1)-\beta^{n-1}(D_2-\alpha D_1)]$.

当 $\alpha=\beta$ 时，$D_n-\alpha D_{n-1}=\alpha(D_{n-1}-\alpha D_{n-2})=\cdots=\alpha^{n-2}(D_2-\alpha D_1)$，

$$\begin{aligned}D_n&=\alpha D_{n-1}+\alpha^{n-2}(D_2-\alpha D_1)\\&=\alpha(\alpha D_{n-2}-\alpha^{n-3}(D_2-\alpha D_1))+\alpha^{n-2}(D_2-D_1)\\&=\alpha^2 D_{n-2}+2\alpha^{n-2}(D_2-\alpha D_1)=\cdots\\&=\alpha^{n-1}D_1+(n-1)\alpha^{n-2}(D_2-\alpha D_1).\end{aligned}$$

例 2.3.5　计算行列式

$$D_n=\begin{vmatrix}a&b&\cdots&b\\c&a&&\vdots\\\vdots&&\ddots&b\\c&\cdots&c&a\end{vmatrix}.$$

解　1）如果 $b=c$，各行加到第一行上，在第一行提取$[a+(n-1)b]$，再用第一行的 $-b$ 倍加到各行得

$$D_n=[a+(n-1)b]\begin{vmatrix}1&1&\cdots&1\\0&a-b&&\vdots\\\vdots&&\ddots&0\\0&\cdots&0&a-b\end{vmatrix}=[a+(n-1)b](a-b)^{n-1}.$$

2）如果 $b\neq c$，将最后一行中的 a 写成$(a-b)+b$ 或$(a-c)+c$ 形式，并分别写成两个行列式的和形式，于是分别得到两个递推公式

$$D_n=(a-b)D_{n-1}+c(a-b)^{n-1};$$

$$D_n=(a-c)D_{n-1}+b(a-c)^{n-1}.$$

将以上两式联立解之得 $D_n=[b(a-c)^n-c(a-b)^n]/(b-c)$.

5. 数学归纳法

例 2.3.6　计算行列式 $D_n=\begin{vmatrix}a_1+x_1&a_2&a_3&\cdots&a_n\\-x_1&x_2&&&\\&-x_2&x_3&&\\&&\ddots&\ddots&\\&&&-x_{n-1}&x_n\end{vmatrix}$

解　$D_1=a_1+x_1=x_1(1+a_1/x_1)$；

$$D_2=\begin{vmatrix}a_1+x_1&a_2\\-x_1&x_2\end{vmatrix}=x_2(a_1+x_1)+a_2x_1=x_1x_2(1+a_1/x_1+a_2/x_2);$$

由此推测 $D_n = x_1x_2\cdots x_n(1+a_1/x_1+a_2/x_2+\cdots+a_n/x_n)$，

事实上，
$$\begin{aligned}D_n &= x_{n-1}D_{n-1}+(-1)^{n-1}a_n(-x_1)(-x_2)\cdots(-x_{n-1})\\ &= x_1x_2\cdots x_n(1+a_1/x_1+a_2/x_2+\cdots+a_{n-1}/x_{n-1})+x_1x_2\cdots x_{n-1}a_n\\ &= x_1x_2\cdots x_n(1+a_1/x_1+a_2/x_2+\cdots+a_n/x_n),\end{aligned}$$

即

$$D_n = \prod_{i=1}^{n} x_i\left(1+\sum_{j=1}^{n}\frac{a_j}{x_j}\right).$$

6. 升阶法

把 n 阶行列式适当地添加 m 行，m 列（$m\geqslant 1$）使得到的 $n+m$ 阶行列式与原行列式的值不变，而且这个升阶后的行列式易于计算，进而求出原行列式的值.

例 2.3.7　计算 n 阶行列式

$$D_n = \begin{vmatrix} a_1+b_1 & a_2 & \cdots & a_n \\ a_1 & a_2+b_2 & & \vdots \\ \vdots & & \ddots & a_n \\ a_1 & \cdots & a_{n-1} & a_n+b_n \end{vmatrix}，其中\ b_1b_2\cdots b_n\neq 0.$$

解　加边为

$$\begin{vmatrix} 1 & a_1 & a_2 & \cdots & a_n \\ 0 & a_1+b_1 & a_2 & \cdots & a_n \\ \vdots & a_1 & a_2+b_2 & \cdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & a_n \\ 0 & a_1 & a_2 & \cdots & a_n+b_n \end{vmatrix}$$

$$\xlongequal{-1(1)+(i)} \begin{vmatrix} 1 & a_1 & \cdots & a_n \\ -1 & b_1 & & 0 \\ \vdots & & \ddots & \vdots \\ -1 & 0 & \cdots & b_n \end{vmatrix}$$

$$\xlongequal{-\frac{1}{b_j}(j+1)+(1)} b_1b_2\cdots b_n\left(1+\sum_{j=1}^{n}\frac{a_j}{b_j}\right).$$

7. 换元法

用同一个元素 x 加到 n 阶行列式 D 中的每一元素上得到 $\overline{D} = D + x\sum\limits_{i,j=1}^{n}A_{ij}$.

一般地，用 $x_1, x_2 \cdots x_n$ 分别加到 n 阶行列式 D 中第 $1,2 \cdots n$ 列的每一元素上得到一个新 n 阶行列式 $\overline{D}$，那么 $\overline{D} = D + \sum_{j=1}^{n} x_j \sum_{i=1}^{n} A_{ij}$.

例 2.3.8 计算行列式 $D_n = \begin{vmatrix} a_1 & x & \cdots & x \\ x & a_2 & & \vdots \\ \vdots & & \ddots & x \\ x & \cdots & x & a_n \end{vmatrix}$.

解 把 D_n 看做是 $\Delta_n = \begin{vmatrix} a_1 - x & 0 & \cdots & 0 \\ 0 & a_2 - x & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & a_n - x \end{vmatrix}$ 中每个元素加上 x 所得，

则
$$D_n = \Delta_n + x\sum_{i=1}^{n}\sum_{j=1}^{n} A_{ij} = \prod_{i=1}^{n}(a_i - x) + x\sum_{j=1}^{n} \frac{\prod_{i=1}^{n}(a_i - x)}{a_j - x}$$
$$= x\prod_{i=1}^{n}(a_i - x)\left(\frac{1}{x} + \sum_{j=1}^{n}\frac{1}{a_j - x}\right).$$

8. *应用典型行列式法*

此法通常是结合前面诸法将行列式变形，化为典型行列式.

例 2.3.9 计算

$$D_n = \begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 2 & 2^2 & 2^3 & \cdots & 2^n \\ 3 & 3^2 & 3^3 & \cdots & 3^n \\ \vdots & \vdots & \vdots & & \vdots \\ n & n^2 & n^3 & \cdots & n^n \end{vmatrix}$$

解 由 $1,2 \cdots n$ 行分别提取 $1,2 \cdots n$，得

$$D_n = n!\begin{vmatrix} 1 & 1 & 1 & \cdots & 1 \\ 1 & 2 & 2^2 & \cdots & 2^{n-1} \\ 1 & 3 & 3^2 & \cdots & 3^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & n & n^2 & \cdots & n^{n-1} \end{vmatrix}$$

再由范德蒙行列式即得

$$D_n = n!(n-1)!(n-2)!\cdots 2!1!$$
$$= n(n-1)^2(n-2)^3\cdots 2^{n-1}.$$

2.4　线性方程组理论

本节主要讨论矩阵的初等变换、矩阵的秩、线性方程组解的存在性、齐次线性方程组的基础解系等问题.

2.4.1　矩阵的初等变换和矩阵的秩

定义 2.4.1　矩阵中不为零子式的最大阶数,称为**矩阵的秩**,记为:秩 A 或 $R(A)$.

显然,设 $A=(a_{ij})\in M_{mn}(F)$,则 $R(A)\leqslant \min\{m, n\}$. 设 A 为 m 行 n 列矩阵 $(a_{ij})_{m\times n}$,

1) 交换行列式的两行(列),用 (i,j)、$[i,j]$ 表示;

2) 用一个数乘矩阵的某一行(列),用 $k(i)$、$k[i]$ 表示;

3) 用第 j 行(列) 的 k 倍加到第 i 行(列),用 $k(j)+(i)$ $k[j]+[i]$ 表示.

定理 2.4.1　初等变换不改变矩阵的秩.

定理 2.4.2　矩阵 $A=(a_{ij})_{m\times n}$ 经初等变换可化成 $\begin{pmatrix} I_r & C_{r\times(n-r)} \\ 0 & 0 \end{pmatrix}$,$r$ 为矩阵A 的秩.

定理 2.4.3　设 $A=(a_{ij})_{m\times n}$,$B=(b_{ij})_{n\times p}$,则 $R(AB)\leqslant \min\{R(A), R(B)\}$.

特别,当其中有一逆阵时,乘积的秩等于另一矩阵的秩.

定理 2.4.4(线性方程组解的存在性定理)　对于线性方程组 $A_{m\times n}X=B_{m\times 1}$,其中 $B=(b_1,b_2,\cdots,b_m)'$. 系数矩阵 A,增广矩阵$\overline{A}$,当秩 $A=$ 秩$\overline{A}=r$ 时,线性方程组有解,当 $r=n$ 时,线性方程组有唯一解;当 $r<n$ 时,线性方程组有无穷多解.

例 2.4.1　讨论以下 $n(\geqslant 2)$ 阶方阵 A 的秩:

$$A=\begin{pmatrix} a & b & \cdots & b \\ b & a & & \vdots \\ \vdots & & \ddots & b \\ b & \cdots & b & a \end{pmatrix}.$$

解　对矩阵 A 作初等变换化为上三角形矩阵:

$$A\rightarrow\begin{pmatrix} a+(n-1)b & b & \cdots & b \\ 0 & a-b & & \vdots \\ \vdots & & \ddots & b \\ 0 & \cdots & 0 & a-b \end{pmatrix},$$

当 $a \neq b$ 且 $a \neq -(n-1)b$ 时，$R(A)=n$；

当 $a=b=0$ 时，$R(A)=0$；

当 $a=-(n-1)b$ 时，$R(A)=n-1$.

2.4.2 向量的线性相关性

定义 2.4.2 向量 α 为向量组$\{\beta_1, \cdots, \beta_s\}$的一个**线性组合**指的是数域 F 中有数 $k_1,\cdots,k_s$，使 $\alpha=k_1\beta_1+\cdots+k_s\beta_s$，此时，又称 α 可由 $\beta_1, \cdots, \beta_s$ **线性表示**.

定义 2.4.3 向量组$\{\alpha_1,\cdots,\alpha_s\}$ $(s\geqslant 1)$ 称为**线性相关**，若数域 F 中有一组不全为零的数 $k_1,\cdots,k_s$，使 $k_1\alpha_1+\cdots+k_s\alpha_s=0$ 成立.

定义 2.4.4 两个向量组$\{\alpha_1,\cdots,\alpha_s\}$与$\{\beta_1, \cdots, \beta_s\}$ **等价**指的是它们中每一组的每一向量均可由另一组的向量线性表示.

定义 2.4.5 向量组的一个部分向量组称为一个**极大线性无关组**，假如这个部分向量组是线性无关的，并且向量组中每一向量均可由这一部分向量组线性表示.

定义 2.4.6 **向量组的秩**指的是它的极大线性无关组所含向量的个数.

定理 2.4.5 矩阵的行向量组的秩与列向量组的秩相等.

定义 2.4.7(矩阵的秩的另一个定义) 把一个矩阵的行向量组的秩(称行秩)与列向量组的秩(即列秩)统称为矩阵的秩.

2.4.3 向量线性相关的性质

定理 2.4.6 若一向量组中一部分向量线性相关，则这一向量组就线性相关. 换言之，若一向量组线性无关，则它的任一非空的部分向量组也线性无关.

定理 2.4.7 设 β 可由 $\alpha_1,\cdots,\alpha_r$ 线性表示，那么向量组$\{\alpha_1,\cdots,\alpha_r\}$线性无关当且仅当 β 可由$\{\alpha_1,\cdots,\alpha_r\}$唯一地线性表示.

定理 2.4.8 $\{\alpha_1,\cdots,\alpha_r\}$线性相关 $\Leftrightarrow$ 其中某一向量是其余向量的线性组合.

定理 2.4.9 向量组中的任何一个线性无关组都可以扩充为它的极大无关组.

定理 2.4.10(替换定理) 设$\{\alpha_1,\cdots,\alpha_r\}$线性无关，$\{\beta_1, \cdots, \beta_s\}$为一个向量组，且$\{\alpha_1,\cdots,\alpha_r\}$可由$\{\beta_1, \cdots, \beta_s\}$线性表示，则 $r\leqslant s$；并且适当调整$\{\beta_1, \cdots, \beta_s\}$的次序，使得用 $\alpha_1,\cdots,\alpha_r$ 替换 $\beta_1, \cdots, \beta_s$ 后，所得的向量组$\{\alpha_1,\cdots,\alpha_r\}$与$\{\beta_1, \cdots, \beta_s\}$等价.

等价命题：若$\{\alpha_1,\cdots, \alpha_r\}$可由$\{\beta_1, \cdots, \beta_s\}$线性表示，且 $r>s$，则$\{\alpha_1,\cdots, \alpha_r\}$必线性相关.

两个等价的线性无关的向量组所含的向量个数必相等，进而等价的向量组必有相同的秩.

特别地，一个向量组的任意两个极大无关组含有相同个数的向量. n 维向量空间中

任意 $n+1$ 个向量必线性相关.

例 2.4.2　设 $A=(a_{ij})_n \in M_n(\mathbf{R})$，如果 $|a_{ii}| > \sum_{i\neq j} |a_{ij}|, j=1,\cdots,n$，那么 $|A| \neq 0$.

证明　设 $\alpha_i=(a_{1i},\cdots,a_{ni})', i=1,\cdots,n$. 如果 $\alpha_1,\cdots,\alpha_n$ 线性无关，则有 $|A|\neq 0$. 事实上，若 $\alpha_1,\cdots,\alpha_n$ 线性相关，则存在一组不全为零的数：$k_1,\cdots,k_n$，使 $\sum_1^n k_i\alpha_i=0$，令 $k=\max\{|k_1|,\cdots,|k_n|\}$，不妨设 $k=|k_1|$，那么 $\alpha_i=\sum_{i\neq j}\left(-\frac{k_j}{k_i}\right)\alpha_j$，特别地 $a_{ii}=\sum_{i\neq j}\left(-\frac{k_j}{k_i}\right)a_{ij}$，则 $|a_{ij}| \leqslant \sum_{i\neq j}\left(-\frac{k_j}{k_i}\right)|a_{ij}| \leqslant \sum_{i\neq j}|a_{ij}|$，这与假设矛盾，故 $\alpha_1,\cdots,\alpha_n$ 线性无关，即 $|A|\neq 0$.

2.4.4　线性方程组有解的判别方法

定义 2.4.8　齐次线性方程组 $A_{mn}X=0$，即

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=0 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=0 \\ \cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=0 \end{cases} \tag{2.4}$$

的一组解向量 $\eta_1,\eta_2,\cdots,\eta_t$ 称为它的一个**基础解系**，如果 $\eta_1,\eta_2,\cdots,\eta_t$ 线性无关，那么这个齐次线性方程组的任一解向量都能表成 $\eta_1,\eta_2,\cdots,\eta_t$ 的线性组合. 把齐次线性方程组 $A_{mn}X=0$ 的一个基础解系生成的空间称为 $AX=0$ 的**解空间**.

定义 2.4.9　通常把齐次线性方程组(2.4) 称为一般线性方程组(2.5) 的**导出方程组**：

$$\begin{cases} a_{11}x_1+a_{12}x_2+\cdots+a_{1n}x_n=b_1 \\ a_{21}x_1+a_{22}x_2+\cdots+a_{2n}x_n=b_2 \\ \cdots\cdots\cdots \\ a_{m1}x_1+a_{m2}x_2+\cdots+a_{mn}x_n=b_m \end{cases} \tag{2.5}$$

定理 2.4.11　设非齐次线性方程组(2.5) 的导出齐次线性方程组为(2.4)，则

1) 非齐次线性方程组(2.5) 有解 $\Leftrightarrow$ 系数阵与增广阵有相同的秩.

2) 方程组(2.5) 若有解，则 ① 当 $R(A)=r<n$ 时，非齐次线性方程组(2.5) 有无穷多个解；② 当 $R(A)=r=n$ 时，非齐次线性方程组(2.5) 有唯一解.

3) 齐次线性方程组(2.4) 有非零解 $\Leftrightarrow$ 它的系数阵的秩小于未知量的个数，即 $R(A)<n$.

4) 当 $m=n$ 时，齐次线性方程组有非零解的充要条件是它的系数行列式为零.

5) 任何一个有非零解的齐次线性方程组必有基础解系，并且基础解系所含解的个数为 $n-R(A)$，即为自由未知量的个数.

6) 若非齐次线性方程组(2.5)有解，则(2.5)的一个解与它的导出组(2.4)的一个解的和是(2.5)的一个解.

例 2.4.3　a 为何值时，以下线性方程组有解，并写出一般解.

$$\begin{cases}3ax_1+(2a+1)x_2+(a+1)x_3=a\\(2a-1)x_1+(2a-1)x_2+(a-2)x_3=a+1\\(4a-1)x_1+3ax_2+2ax_3=1\end{cases}$$

解　系数行列式 $D=|A|=\begin{vmatrix}3a & 2a+1 & a+1\\2a-1 & 2a-1 & a-2\\4a-1 & 3a & 2a\end{vmatrix}=(a-1)^2(a+1)$，所以

(1) 当 $a\neq\pm 1$ 时，$D\neq 0$，由克莱姆规则，方程组有唯一解：

$$x_1=\frac{4a+1}{(a-1)(a+1)},x_2=\frac{a(2a-7)}{(a-1)(a+1)},x_3=\frac{-3a}{a+1}.$$

(2) 当 $a=1$ 时，方程组有无穷多解：$x_1=1-k,x_2=k,x_3=-1,k\in F$；

(3) 当 $a=-1$ 时，方程组无解.

2.4.5　矩阵的秩与向量的线性相关性

定理 2.4.12　齐次线性方程组 $AX=0$ 的解都是 $BX=0$ 的解，则 $R(A)\geqslant R(B)$.

证明　设 $AX=0$ 的基础解系为 $\xi_1,\cdots,\xi_{n-r}$；$BX=0$ 的基础解系为 $\eta_1,\cdots,\eta_{n-s}$，其中 $r=R(A)$，$s=R(B)$，则由题设知 $\xi_1,\cdots,\xi_{n-r}$ 可由 $\eta_1,\cdots,\eta_{n-s}$ 线性表出，由替换定理，必有 $n-r\leqslant n-s$，即 $r\geqslant s$，亦即 $R(A)\geqslant R(B)$.

定理 2.4.13(秩的第一降阶定理)　设 A 为 n 阶可逆阵，D 为 m 阶阵，B，C 分别为 $n\times m$，$m\times n$ 阵，则 $R\begin{pmatrix}A & B\\C & D\end{pmatrix}=R(A)+R(D-CA^{-1}B)$.

证明　由 A 可逆，则有 $\begin{pmatrix}I_n & 0\\-CA^{-1} & I_m\end{pmatrix}\begin{pmatrix}A & B\\C & D\end{pmatrix}=\begin{pmatrix}A & B\\0 & D-CA^{-1}B\end{pmatrix}$，

故 $R\begin{pmatrix}A & B\\C & D\end{pmatrix}=R\begin{pmatrix}A & B\\0 & D-CA^{-1}B\end{pmatrix}=R(A)+R(D-CA^{-1}B)$.

由降阶定理，又可得到秩的升阶公式：

$$R(D-CA^{-1}B)=R\begin{pmatrix}A & B\\ C & D\end{pmatrix}-R(A).$$

再由此来证明 Sylvester 不等式是很简单的，

$$R(AB)=R[0-(-A)I_nB]=R\begin{pmatrix}I_n & B\\ -A & 0\end{pmatrix}-R(I_n)=R\begin{pmatrix}B & I_n\\ 0 & -A\end{pmatrix}-n$$

$$\geqslant R(B)+R(-A)-n=R(A)+R(B)-n.$$

当 D 可逆时，秩的第一降阶定理为 $R\begin{pmatrix}A & B\\ C & D\end{pmatrix}=R(D)+R(A-BD^{-1}C).$

特别地，当 A,D 可逆时，$R(A)-R(D)=R(A-BD^{-1}C)-R(D-CA^{-1}B)$，称为第二降阶定理.

例 2.4.4　设系数阵的秩为 r 的含有 n 个未知量的非齐次线性方程组 $AX=B$ 有解. 试证，它的任一解（向量）都可由它的 $n-r+1$ 个线性无关的解向量线性表示.

证明　设 η_0 为 $AX=B$ 的一个特解；$\eta_1,\cdots,\eta_{n-r}$ 为它的导出方程组 $AX=0$ 的一个基础解系，于是 $AX=B$ 的任一解（向量）为

$$\eta=\eta_0+\sum_{i=1}^{n-r}k_i\eta_i=(1-\sum_{i=1}^{n-r}k_i)\eta_0+\sum_{i=1}^{n-r}k_i(\eta_0+\eta_i).$$

令 $k_0=1-\sum\limits_{i=1}^{n-r}k_i$，则 $\eta=k_0\eta_0+\sum\limits_{i=1}^{n-r}k_i(\eta_0+\eta_i)$，其中 η_0，$\eta_0+\eta_1$，$\eta_0+\eta_2$，$\cdots$，$\eta_0+\eta_{n-r}$ 显然是 $AX=B$ 的 $n-r+1$ 个解向量，这 $n-r+1$ 个解向量是线性无关的.

若 $x_0\eta_0+\sum\limits_{i=1}^{n-r}x_i(\eta_0+\eta_i)=0$，则 $(\sum\limits_{j=0}^{n-r}x_j)\eta_0=-\sum\limits_{i=1}^{n-r}x_i\eta_i$，对上式两端同左乘 A 得

$$(\sum_{j=0}^{n-r}x_j)A\eta_0=-\sum_{i=1}^{n-r}x_iA\eta_i.$$

由 $A\eta_0=B,A\eta_i=0,i=1,\cdots,n-r$，有 $\sum\limits_{j=0}^{n-r}x_jB=0$. 因 $B\neq0$，推出 $\sum\limits_{j=0}^{n-r}x_j=0$. 把它代入上式左端又得 $\sum\limits_{j=0}^{n-r}x_j\eta_j=0$，则由 $\eta_1,\cdots,\eta_{n-r}$ 线性无关推得 $x_1=\cdots=x_{n-r}=0$，进而 $x_0=0$. 所以 $AX=B$ 的任一解向量 η 均可由 $n-r+1$ 个线性无关的解向量 η_0，$\eta_0+\eta_1$，$\eta_0+\eta_2$，$\cdots$，$\eta_0+\eta_{n-r}$ 线性表示.

2.5　矩阵代数理论

矩阵论是线性代数的主要研究对象之一，矩阵又是处理线性代数问题的主要工具，

因此,必须熟练地掌握矩阵理论中解决各种问题的思想方法与途径,这是极为重要的.矩阵的初等变换是矩阵论中最基本的内容,是研讨矩阵有关问题的重要方法,特别是分块矩阵的初等变换,即广义初等变换,更是原理简明,内容具体,表述清晰,易于掌握与使用,在论述矩阵分解、矩阵可逆性的判别、矩阵方程、矩阵的化简及标准形等诸多问题中都有广泛的应用.

2.5.1 矩阵的运算

定义 2.5.1(矩阵的加法) 两个 m 行 n 列矩阵 $A=(a_{ij})_{m\times n}$, $B=(b_{ij})_{m\times n}$ 的加法指的是 $A+B=(a_{ij}+b_{ij})_{m\times n}$.

(数与矩阵的乘法) $k\in F$, k 与矩阵 $A=(a_{ij})_{m\times n}$ 之乘法指的是 $kA=(ka_{ij})_{m\times n}$.

(矩阵的乘法) m 行 n 列矩阵 $A=(a_{ij})_{m\times n}$ 与 n 行 p 列矩阵 $B=(b_{ij})_{n\times p}$ 的乘法指的是 $AB=(c_{ij})_{m\times p}$,其中 $c_{ij}=a_{i1}b_{1j}+a_{i2}b_{2j}+\cdots+a_{in}b_{nj}$.

定理 2.5.1 1) 矩阵乘积的行列式等于它的因子的行列式乘积,即 $|AB|=|A||B|$,其中 A,B 皆为 n 阶矩阵.

2) 矩阵乘积的秩不超过其因子的秩,即 $R(AB)\leqslant\min\{R(A),R(B)\}$.

特别地,当其中有因子为可逆时,乘积之秩等于另一因子之秩.

定义 2.5.2 **矩阵的转置**指的是将矩阵 $A=(a_{ij})_{m\times n}$ 的每一行变为相应的列所得到的另一矩阵,记作 $A'=(a_{ji})_{n\times m}$.

转置矩阵的性质:

1) $(A')'=A$;

2) $(A+B)'=A'+B'$, $(A_1+A_2+\cdots+A_r)'=A'_1+A'_2+\cdots+A'_r$;

3) $(AB)'=B'A'$, $(A_1A_2\cdots A_r)'=A'_r\cdots A'_2A'_1$;

4) $(aA)'=aA'$.

例 2.5.1 任何一个 n 阶矩阵 A,均可表为一个对称阵与一个反对称阵之和,即 $A=B+C$,其中 $B'=B$, $C'=-C$.

证明 对于 $A=B+C$,等号两边取转置得: $A'=B'-C'$,把两式联立起来解得:

$$B=A/2+A'/2, C=A/2-A'/2,$$

易验证 $A=B+C$,且 $B'=B$, $C'=-C$,即 B 为对称阵,C 为反对称阵.

2.5.2 初等矩阵

定义 2.5.3 由单位矩阵 I 经过一次初等变换得到的矩阵称为**初等矩阵**.于是初等矩阵有

(1) 换法矩阵 P_{ij}；　(2) 倍法矩阵 $D_i(k)$；　(3) 消法矩阵 $T_{ij}(k)$；

这三种初等矩阵都是由单位矩阵经过初等变换而得到的矩阵.

每一初等矩阵都与一个初等变换相对应. 初等矩阵的作用在于将矩阵的初等变换的"几何"语言转换为矩阵乘法的"代数"语言，即矩阵的行的初等变换表述为它的左侧乘以相应初等矩阵，而列的初等变换则是它的右侧乘以相应初等矩阵. 这种表述方式，简明、严谨，充分体现了数学美.

定理 2.5.2　任意一个 $m\times n$ 矩阵 A，都存在一个 m 阶可逆阵 P，一个 n 阶可逆阵 Q，使 $PAQ=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$，其中 I_r 为 r 阶单位矩阵，而 $r=R(A)$.

定理 2.5.3　设 $m\times n$ 矩阵 A 的秩是 r，且 $PAQ=\begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$，其中 P 为 m 阶可逆阵，Q 为 n 阶可逆阵，则 A 的全部广义逆为 $G=Q\begin{pmatrix} I_r & C \\ D & F \end{pmatrix}P$. 其中 C,D,F 分别为 $r\times(m-n)$，$(n-r)\times r$，$(n-r)\times(m-r)$ 阶矩阵.

此结论说明任何矩阵的广义逆都存在，并且不唯一.

定义 2.5.4　矩阵的分块指的是在它的行或列之间加上一些线，把这一矩阵分成若干小块. 而用这种方法分成若干小块的矩阵称为一个**分块矩阵**.

1) 两个同型矩阵可以相加，只需把对应位置的小块相加；

2) 数乘分块矩阵，只需用这个数乘遍每一小块；

3) 两个可乘矩阵 A、B，若 A 的列的分法与 B 的行的分法一致，则 A 与 B 可以相乘，只需按数字矩阵乘法规则，把小矩阵当作数一样来处理.

4) 广义初等变换与广义初等矩阵：

把分块矩阵中每一小块看作一个数，类似于数字矩阵的三种初等变换：

Ⅰ) 交换两行(列)；

Ⅱ) 某一行(列) 左(右) 乘一个非零矩阵 P；

Ⅲ) 某一行(列) 左(右) 乘以矩阵 P 加到另一行(列) 对应元素上去，统称为分块矩阵的初等变换，简称为广义初等变换，得到的矩阵，叫广义初等矩阵.

2.5.3　可逆矩阵及其性质

定义 2.5.5　已知 $A\in M_n(F)$，若存在矩阵 $B\in M_n(F)$，使得 $AB=BA=I$，则称 A 为一个**可逆矩阵**(或**非奇异矩阵**)，而 B 称为 A 的**逆阵**；否则称 A 为**不可逆矩阵**.

定理 2.5.4　设 A 为 n 阶可逆矩阵，则当且仅当是下列条件之一：

1) 存在 $B\in M_n(F)$，使 $AB=BA=I$；

2) A 经初等变换可化为 n 阶单位矩阵 I；

3) A 可表为若干初等矩阵之积，即 $A = E_1E_2\cdots E_s$；

4) 存在 n 阶可逆矩阵 P,Q，使 $PAQ = I$；

5) A 的秩为 n，即 $R(A) = n$；

6) A 的行列式非零，即 $|A| \neq 0$；

可逆矩阵的性质如下：

1) 可逆矩阵 A 的逆阵是唯一的，记为 A^{-1}；

2) 可逆矩阵 A 的逆阵 A^{-1} 也是可逆的，并且 $(A^{-1})^{-1} = A$；

3) 两个同阶可逆矩阵 A 与 B 的乘积 AB 也可逆，并且 $(AB)^{-1} = B^{-1}A^{-1}$，若 $A_1,\cdots,A_r$ 为同阶可逆矩阵 $(r \geqslant 2)$，则 $A_1A_2\cdots A_r$ 也可逆，且

$$(A_1A_2\cdots A_r)^{-1} = A_r^{-1}\cdots A_2^{-1}A_1^{-1}.$$

4) 若 A 为可逆矩阵，则 $-A, A', \overline{A}, kA\,(k \neq 0)$ 均可逆，它们的逆分别是：

$$(-A)^{-1} = -A^{-1};\ (A')^{-1} = (A^{-1})';\ (\overline{A})^{-1} = \overline{A^{-1}};\ (kA)^{-1} = k^{-1}A^{-1}.$$

其中，$\overline{A}$ 表示 A 的共轭，即 A 的每个元素（数）的共轭（复数），按照 A 的原来的相对位置排成的同阶矩阵.

5) 初等矩阵都是可逆的，它们的逆矩阵还是初等矩阵，即

$$P_{ij}^{-1} = P_{ij};\quad D_i^{-1}(k) = D_i(1/k);\quad T_{ij}^{-1}(k) = T_{ij}(-k).$$

定义 2.5.6　n 阶矩阵 A 的**伴随矩阵** A^* 指的是由矩阵 A 的所有元素 $a_{ij}(i,j = 1,\cdots,n)$ 的代数余子式 A_{ij} 作为元素所构成的同阶矩阵.

矩阵 A 的伴随矩阵的性质：

1) $AA^* = A^*A = |A|I$；$|A^*| = |A|^{n-1}$；

2) $(A^*)' = (A')^*$；

3) $(aA)^* = a^{n-1}A^*$；$(A^*)^* = |A|^{n-1}A$；$(AB)^* = B^*A^*$

4) 若 A 可逆，则 $A^* = |A|A^{-1}$，且 A^* 可逆，其逆为 $(A^*)^{-1} = (A^{-1})^* = |A|^{-1}A$；

5) $R(A^*) = \begin{cases} n \Leftrightarrow R(A) = n \\ 1 \Leftrightarrow R(A) = n-1 \\ 0 \Leftrightarrow R(A) < n-1 \end{cases}$

例 2.5.2　设 A 与 B 分别是 $n\times m$ 阵与 $m\times n$ 阵，则 $I_m - BA$ 可逆 $\Leftrightarrow I_n - AB$ 可逆.

证明　$A(I_m - BA) = (I_n - AB)A$，若 $I_m - BA$ 可逆，则 $A = (I_n - AB)A(I_m - BA)^{-1}$，于是

$$I_n = (I_n - AB) + AB = (I_n - AB) + (I_n - AB)A(I_m - BA)^{-1}B$$

$= (I_n - AB)[I_n + A(I_m - BA)^{-1}B]$,

故$(I_n - AB)$可逆,且

$$(I_n - AB)^{-1} = I_n + A(I_m - BA)^{-1}B.$$

同理由 $I_n - AB$ 可逆,可证 $I_m - BA$ 可逆,且

$$(I_m - BA)^{-1} = I_m + B(I_n - AB)^{-1}A.$$

定义 2.5.7　设 A 为 $m \times n$ 矩阵,若 $n \times m$ 矩阵 G 满足 $AGA = A$,则称 G 为 A 的一个**广义逆矩阵**,简称**广义逆**.

特别地,当 A 为 n 阶可逆阵时,由 $AGA = A$ 即得 $G = A$,此时广义逆即为逆阵,故广义逆矩阵是通常逆矩阵概念的推广. 这种广义逆的定义,决定了矩阵的广义逆是不唯一的,这对某些问题的研究带来不便,在实际问题中有时要求有特殊条件的广义逆. 这就是 Moore 与 Penrose 给出的

定义 2.5.8(Moore - Penrose 广义逆)　设 A 为 $m \times n$ 复(实)矩阵,若 $n \times m$ 矩阵 G 满足,

(1) $AGA = A$;

(2) $GAG = G$;

(3) $(\overline{AG})' = AG$(AG 为实对称阵);

(4) $(\overline{GA})' = GA$(GA 为实对称阵),

则称 G 为 A 的 **Moore - Penrose 广义逆**,简称 G 为 A 的 **M - P 广义逆**.

上述定义的广义逆是唯一的. 通常 A 的广义逆记为 A^+.

2.5.4　矩阵的特征根、特征向量、特征多项式

定义 2.5.9　设 A 为 n 阶阵,若有一数 λ 及一 n 维非零列向量 X,使得 $AX = \lambda X$,即 $(\lambda I_n - A)X = 0$,则称 λ 为 A 的**特征根**(或称特征值),称非零向量 X 为 A 的属于特征根 λ 的**特征向量**,又称 $|\lambda I_n - A| = f_A(\lambda)$ 为 A 的**特征多项式**.

定义 2.5.10　设 A 为 n 阶方阵,$f(x)$ 为数域 F 上一个 m 次多项式,则称 $f(A)$ 为 A 的**矩阵多项式**,并且也是一个 n 阶方阵;

若 $f(A) = 0$,则称 $f(x)$ 为 A 的**化零多项式**;

称 A 的所有化零多项式中次数最小的首一的多项式为 A 的**最小多项式**.

定理 2.5.5　设 $A = (a_{ij})_n$,则 A 的特征多项式

$$f_A(\lambda) = |\lambda I_n - A| = \lambda^n - \mathrm{Tr}(A)\lambda^{n-1} + \cdots + (-1)^n |A|,$$

其中 $\mathrm{Tr}(A)$ 为 A 的主角线上 n 个元素之和,称矩阵 A 的**迹**,并且 $\mathrm{Tr}(A)$ 等于 A 的全部

特征根的和，$\det A$ 等于 A 的全部特征根的积.

定理 2.5.6(Hamilton - Cayley 定理) 设 $f(\lambda)$ 为 n 阶矩阵 A 的特征多项式，则 $f(A)=0$，即 A 的特征多项式是它的化零多项式.

定理 2.5.7 设 A 为 n 阶正定矩阵，则下列命题等价：

(1) 存在 n 阶正交矩阵 U，使 $U'AU=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$，其中 $\lambda_i>0(i=1,\cdots,n)$；

(2) A 的所有顺序主子式全大于零；

(3) 对 $\forall n$ 维非零实列向量 X，都有 $X'AX>0$.

矩阵最小多项式的性质如下：

(1) 方阵 A 的最小多项式 $m_A(\lambda)$ 必整除 A 的特征多项式 $f_A(\lambda)$，即 $m_A(\lambda)\mid f_A(\lambda)$；

(2) 方阵 A 的最小多项式是唯一的；

(3) 相似矩阵有相同的最小多项式；

(4) 设 $A=\mathrm{diag}(A_1,A_2,\cdots,A_s)$ 为分块对角阵，A 的最小多项式为 $m(\lambda)$，A_i 的最小多项式为 $m_i(\lambda)$，$i=1\cdots s$，则 $m(\lambda)$ 为 $m_1(\lambda),\cdots,m_s(\lambda)$ 的最小公倍式.

定义 2.5.11 设 A,B 都是 n 阶方阵，若存在一个 n 阶可逆阵 P，使 $P^{-1}AP=B$，则称 A 与 B **相似**，或称 A 相似于 B，记作 $A\sim B$.

特别地，对 n 阶阵 A，若存在 n 阶可逆阵 P，使 $P^{-1}AP=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$，即 A 相似于对角阵，此时称 A **可对角化**，并称对角阵 $\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$ 为 A 的**相似标准形**，其中 $\lambda_1,\cdots,\lambda_n$ 为 A 的全部特征根.

定义 2.5.12 设 A,B 都是 n 阶方阵，若存在一个 n 阶可逆阵 P，使 $B=P'AP$，则称 A 与 B **合同**，或称 A 合同于 B.

例 2.5.3 设 A 为 n 阶可逆实矩阵，则 A 可分解为 $A=P\mathrm{diag}(\lambda_1,\cdots,\lambda_n)T$，其中 P,T 皆为 n 阶正交阵，$\lambda_i>0,i=1,\cdots,n$，且 $\lambda_1^2,\cdots,\lambda_n^2$ 是矩阵 $A'A$ 的全部特征根.

证明 由于 A 是 n 阶实可逆阵，必存在一个 n 阶正交阵 Q，一个 n 阶正定阵 S，使 $A=Q\cdot S$，进而 $A'A=S^2$. 对 n 阶正定阵 S，存在 n 阶正交阵 V，使 $V'SV=\mathrm{diag}(\lambda_1,\cdots,\lambda_n)$，其中 $\lambda_i(i=1,\cdots,n)$ 是正定阵 S 的全部特征根，于是得

$$S=V\mathrm{diag}(\lambda_1,\cdots,\lambda_n)V',S^2=V\mathrm{diag}(\lambda_1^2,\cdots,\lambda_n^2)V'.$$

由于 $S^2=A'A$，即得

$$V'(A'A)V=\mathrm{diag}(\lambda_1^2,\cdots,\lambda_n^2).$$

故 $\lambda_1^2,\cdots,\lambda_n^2$ 是 $A'A$ 的全部特征根，同时，

$$A=Q\cdot S=Q\cdot V\mathrm{diag}(\lambda_1,\cdots,\lambda_n)V'.$$

对此令 $P = QV, T = V'$，则 P, T 皆为 n 阶正交阵，且 $A = P\mathrm{diag}(\lambda_1, \cdots, \lambda_n)T$.

2.5.5　矩阵标准形的初步讨论

1) λ-阵 $A(\lambda)$ 指的是它的元素是关于 λ 的多项式的矩阵；

2) 秩 $A(\lambda) = r$，指的是 $A(\lambda)$ 中有一个 r 阶子式非零，所有 $r+1$ 阶子式皆为零；

3) 若有一个 n 阶 λ-阵 $B(\lambda)$，使 $A(\lambda)B(\lambda) = B(\lambda)A(\lambda) = I_n$，称 $A(\lambda)$ 为可逆的；

4) λ-阵也有三种初等变换；

5) λ-阵 $A(\lambda)$ 与 $B(\lambda)$ 等价指的是 $A(\lambda)$ 经一系列初等变换可化为 $B(\lambda)$.

6) 若任一秩为 r 的非零 n 阶 λ-阵 $A(\lambda)$ 都等价于 λ-阵 $D(\lambda)$，则称具有此性质的对角形 λ-阵 $D(\lambda)$ 为 $A(\lambda)$ 的标准形，标准形是唯一的.

定义 2.5.13　$A(\lambda)$ 的标准形主对角线上非零元素 $d_1(\lambda), \cdots, d_r(\lambda)$ 称为 $A(\lambda)$ 的**不变因子**.

特别地，矩阵 A 的特征矩阵 $\lambda I - A$ 的不变因子称为 A 的不变因子.

矩阵 $A(\lambda)$ 的 k 阶行列式因子指的是 $A(\lambda)$ 中全部 k 阶子式的首一的最大公因式 $D_k(\lambda)$，其中 $1 \leqslant A \leqslant A(\lambda)$ 的秩. 因此，秩为 r 的 λ-阵 $A(\lambda)$ 共有 r 个行列式因子.

n 阶矩阵 A 的**初等因子**指的是把矩阵 A 的每个次数大于 0 的不变因子分解成（在 $\mathbf{C}$ 内）互异的一次因式方幂的乘积（所有这些一次因式方幂，相同的必须按出现的次数计算）.

定义 2.5.14　数域 F 上任何一个 n 阶方阵必相似于一个有理标准形 F，即存在一个 n 阶可逆阵 T，使 $T^{-1}AT = J = \mathrm{diag}(F_1, \cdots, F_s)$，其中 $F_i(i = 1, \cdots, s)$ 为 n_i 阶 Frobenius 块（有理块），$\sum n_i = n$. 通常称 F 为 A 的**有理标准形**（Frobenius 标准形）.

复数域上任何一个 n 阶复阵 A 都与一个 Jordan 形矩阵相似，即存在一个 n 阶可逆复阵 T，使 $T^{-1}AT = J = \mathrm{diag}(J_1, \cdots, J_s)$，其中 $J_i(i = 1, \cdots, s)$ 为 n_i 阶 Jordan 块，$\sum n_i = n$，除去 $J_i(i = 1, \cdots, s)$ 的排列次序外是被 A 唯一决定，通常称 J 为 A 的**若当标准形**（Jordan 标准形）.

定理 2.5.8　一矩阵 $A(\lambda)$ 可逆 $\Leftrightarrow |A(\lambda)| = d \neq 0 \Leftrightarrow A(\lambda)$ 可表为初等 λ-矩阵之积.

定理 2.5.9　λ-阵 $A(\lambda)$ 与 $B(\lambda)$ 等价 $\Leftrightarrow$ 有相同的行列式因子 $\Leftrightarrow$ 有相同的不变因子 $\Leftrightarrow$ 有可逆 λ-阵 $P(\lambda), Q(\lambda)$，使 $B(\lambda) = P(\lambda)A(\lambda)Q(\lambda)$.

定理 2.5.10　数域 F 上 n 阶方阵 A 与 B 相似 $\Leftrightarrow$ 它们的特征矩阵 $\lambda I - A$ 与 $\lambda I - B$ 等价 $\Leftrightarrow$ 它们有相同的不变因子 $\Leftrightarrow$ 有相同的行列式因子 $\Leftrightarrow$ 有相同的初等

因子.

定理 2.5.11(Robenius 定理) 设数域 F 上的 n 阶方阵 A 的特征多项式为 $f(\lambda)$,最小多项式为 $m(\lambda)$,则 $m(\lambda) = d_n(\lambda)$,其中 $d_n(\lambda)$ 为 A 的最后一个不变因子,即 A 的最小多项式就是 A 的最后一个不变因子; $f(\lambda)$ 在数域 **P** 上的任一不可约因式都是 $m(\lambda)$ 的因式.

定理 2.5.12 矩阵 A 可对角化 $\Leftrightarrow A$ 的最小多项式无重根 $\Leftrightarrow A$ 的初等因子全为一次的 $\Leftrightarrow A$ 的不变因子都没有重根.

第3章 向量空间与线性变换

向量空间是最基本的数学概念之一，也是我们接触到的第一个代数系统. 它的理论和方法已经渗透到自然科学、工程技术的各个领域. 本章主要讨论向量空间的概念、子空间的直和分解、向量空间的同构等基本内容. 在向量空间的讨论中，也将加深对线性方程组和矩阵的理解.

在数学里，变换的概念是非常基本的，这里我们主要考虑一个向量空间到自身的线性变换，注意到在有限维条件下，线性变换和矩阵之间有必然的联系，讨论线性变换的对角化及向量空间的准素分解等问题.

3.1 向量空间

3.1.1 向量空间的定义

定义 3.1.1 设 V 是由称为"向量"的元素构成的一个非空集合，F 为一数域，在 V 中定义一个称为"加法"的代数运算，即映射

$$V \times V \to V, \forall \alpha, \beta \in V, \text{使 } \alpha + \beta \in V.$$

在数域 F 与集合 V 的元素间还定义一个称为"数量乘法"的运算，即映射

$$F \times V \to V, \forall a \in F, \alpha \in V, \text{使 } a\alpha \in V.$$

如果加法与数量乘法满足八条公理，那么 V 称为数域 F 上的**向量空间**(或**线性空间**)，记为 $V(F)$，简记为 V.

定义 3.1.2 设 W 为 $V(F)$ 的一个非空子集，若 W 对于 V 的加法与数量乘法来说构成一个向量空间，则称 W 为 V 的一个**子空间**.

定理 3.1.1 设 W 为 $V(F)$ 的一个非空子集，则 W 为 V 的一个子空间的充分必要条件是对 $\forall \alpha, \beta \in W, \forall a, b \in F$，都有 $a\alpha + b\beta \in W$.

定义 3.1.3 设 V_1, V_2 皆为 $V(F)$ 的两个子空间，则称 $V_1 \cap V_2$ 为 V_1 与 V_2 的**交**，是 V 的一个子空间.

定义 3.1.4 设 V_1, V_2 皆为 $V(F)$ 的两个子空间，称

$$V_1 + V_2 = \{\alpha_1 + \alpha_2 \mid \alpha_1 \in V_1, \alpha_2 \in V_2\}$$

为 V_1 与 V_2 的和,是 V 的一个子空间.

定义 3.1.5 设$\{\alpha_1, \alpha_2 \cdots \alpha_n\} \subseteq V(F)$,则称 $L(\alpha_1, \alpha_2 \cdots \alpha_n) = \{k_1\alpha_1 + \cdots + k_n\alpha_n \mid k_i \in F\}(i = 1,2 \cdots n)$ 为由 $\alpha_1, \alpha_2 \cdots \alpha_n$ **生成的子空间**,$L(\alpha_1, \alpha_2 \cdots \alpha_n)$ 还是 $V(F)$ 的一个子空间.

定理 3.1.2 设 $V_1, V_2 \cdots V_s$ 是向量空间 $V(F)$ 的 s 个非平凡子空间,则 V 中至少有一向量 α,使 $\alpha \notin V_i, i = 1,2\cdots s$.

证明 对 s 采用归纳法证之.

(1) 当 $s = 2$ 时,由于 V_1,V_2 为 V 的非平凡子空间,必有 $\alpha_1,\alpha_2 \in V$,使 $\alpha_1 \notin V_1$, $\alpha_2 \notin V_2$.若 $\alpha_1 \notin V_2$,则 α_1 即为所求;

若 $\alpha_2 \notin V_1$,则 α_2 即为所求.

若 $\alpha_1 \in V_2, \alpha_2 \in V_1$,则令 $\alpha = \alpha_1 + \alpha_2$,必有 $\alpha \notin V_i, i = 1,2$,故 $s = 2$ 时结论正确.

(2) 假设 $s-1$ 时命题成立,即存在 $\beta_1 \in V$,使 $\beta_i \notin V_i, i = 1,2\cdots s-1$.若 $\beta_1 \notin V_s$,则结论得证.

设 $\beta_1 \in V_s$,同理 $\exists \beta_2 \in V$,使 $\beta_2 \notin V_j, j = 2, 3\cdots s$.若 $\beta_2 \notin V_1$,则结论得证.设 $\beta_2 \in V_1$,由 β_1,β_2 作 $\alpha_k = \beta_1 + k\beta_2$,其中 k 为正整数.易知 $\alpha_k \notin V_1, \alpha_k \notin V_s$.

同时,$V_2, V_3 \cdots V_{s-1}$ 这 $s-2$ 个非平凡子空间中每一个至多含一个 α_k.否则,当 $k_1 \neq k_2$ 时,若 $\alpha_{k_1}, \alpha_{k_2} \in V_i, i = 2, 3\cdots s-1$,则由 $\alpha_{k_1} = \beta_1 + k_1\beta_2, \alpha_{k_2} = \beta_1 + k_2\beta_2$ 推得 β_1, $\beta_2 \in V_i$,这与 β_1,β_2 的选取相矛盾.

因此,由正整数 k 的不同而得到无限多个 α_k 中至少有一个 α_{k_0},使 $\alpha_{k_0} \notin V_i, i = 1, 2,\cdots s$,故结论成立.

例 3.1.1 设 $V_1, V_2 \cdots V_s$ 是 n 维向量空间 $V_n(F)$ 的任意 s 个非平凡子空间.试证:存在 V 的一个基,使这个基的 n 个基向量均不在 $V_1, V_2 \cdots V_s$ 中.

证明 由于 $V_1, V_2 \cdots V_s$ 为 V 的非平凡子空间,所以存在 $\alpha_1 \in V$,使 $\alpha_1 \notin V_i, i = 1, 2\cdots s$,对此考虑 V 的非平凡子空间:$V_1, V_2 \cdots V_s$.

又由定理 3.1.1 知,存在 $\alpha_2 \in V$,使 $\alpha_2 \notin V_i, i = 1, 2\cdots s, \alpha_2 \notin L(\alpha_1)$,而由 $\alpha_2 \notin L(\alpha_1)$ 知 α_1,α_2 线性无关.于是必存在 $\alpha_3 \in V$,使 $\alpha_3 \notin V_i, i = 1, 2\cdots s, \alpha_3 \notin L(\alpha_1,\alpha_2)$,且 $\alpha_1,\alpha_2,\alpha_3$ 线性无关(否则,$\alpha_1, \alpha_2, \alpha_3$ 线性相关,而 α_1, α_2 线性无关,推得 α_3 为 α_1, α_2 的线性组合,即 $\alpha_3 \in L(\alpha_1, \alpha_2)$,矛盾).

同时 $\alpha_1,\alpha_2,\alpha_3$ 全不属于 $V_1, V_2 \cdots V_s$,这样经过上述 $n-1$ 次证法后,得到 V 的线性无关向量 $\alpha_1, \alpha_2 \cdots \alpha_{n-1}$,使它们全不在 $V_1,V_2 \cdots V_s$ 中,存在 $\alpha_n \in V$,使 $\alpha_n \notin V_i, i = 1, 2\cdots s, \alpha_n \notin L(\alpha_1, \alpha_2 \cdots \alpha_{n-1})$,且 $\alpha_1, \alpha_2 \cdots \alpha_{n-1}$ 线性无关,因此得到 V 的一个基 $\alpha_1, \alpha_2 \cdots \alpha_n$,使其每一基向量 $\alpha_i, i = 1, 2\cdots n$,全不在 $V_1, V_2 \cdots V_s$ 中.

3.1.2　向量空间的基和维数

定义 3.1.6　向量空间$V(F)$中集合S生成的子空间是S中向量的全体线性组合所成集合

$$< S >= \mathrm{span}(S) = \{r_1v_1 + \cdots + r_nv_n \mid r_i \in F, v_i \in V\}.$$

当$S = \{v_1, v_2 \cdots v_n\}$是有限集时，我们就用符号$< v_1, v_2 \cdots v_n >$或$\mathrm{span}\{v_1, v_2 \cdots v_n\}$. 称向量集$S$生成了$V$.

如果$V = \mathrm{span}(S)$，即对于某些标量$r_1, r_1 \cdots r_n$和向量$v_1, v_2 \cdots v_n$，每个向量$v \in V$都能表示成形为$v = r_1v_1 + \cdots + r_nv_n$. 所有向量空间都有生成集，因为整个空间就是一个生成集.

定义 3.1.7　$V(F)$中非空向量集S**线性无关**，如果对于$V(F)$中$v_1, v_2 \cdots v_n$，有

$$r_1v_1 + \cdots + r_nv_n = 0 \Rightarrow r_1 = \cdots = r_n = 0.$$

如果一个向量集不是线性无关，那么它就是**线性相关**.

从定义出发可知，线性无关集的任何非空子集也线性无关.

定义 3.1.8　向量空间$V(F)$的一个**基**指的是V中一组向量$\{\alpha_1, \cdots, \alpha_n\}$满足

1）$\alpha_1, \alpha_2 \cdots \alpha_n$线性无关；

2）V中每一向量都可由$\alpha_1, \alpha_2 \cdots \alpha_n$线性表示.

简言之，V的一个基就是V的一个极大无关组或一组线性无关的生成元.

定义 3.1.9　**维数**指的是$V(F)$的一个基所含向量个数，记为$\dim V$.

特别地，$\dim\{0\} = 0$. n维向量空间记为$V_n(F)$.

定理 3.1.3　$\{\alpha_1, \cdots, \alpha_n\}$为$V_n(F)$的一个基$\Leftrightarrow V$中每一向量$\xi$都可唯一地表成这$n$个向量的线性组合.

定理 3.1.4　数域F上n维向量空间$V_n(F)$中任意多于n个向量必线性相关. 因此有以下四个重要结论：

1）$V_n(F)$中任意n个线性无关的向量均可构成一个基；

2）$V_n(F)$中任何两个基所含向量个数相同；

3）有限维向量空间的任一子空间必为有限维的；

4）设W_1, W_2为$V_n(F)$的两个子空间，且有$W_1 \subseteq W_2$，若$\dim W_1 = \dim W_2$，则$W_1 = W_2$.

定理 3.1.5(维数定理)　设V_1, V_2皆为$V(F)$的两个有限维子空间，则$V_1 + V_2$也是V的有限维子空间，并且

$$\dim(V_1 + V_2) = \dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2).$$

定义 3.1.10　设$\{\alpha_1, \cdots, \alpha_n\}$与$\{\beta_1, \cdots, \beta_m\}$是向量空间$V(F)$的两个向量组，如果每一个$\alpha_i$可由$\{\beta_1, \cdots, \beta_m\}$线性表示，每一个$\beta_j$可由$\{\alpha_1, \cdots, \alpha_n\}$线性表示，称这两个

向量组等价.

定理 3.1.6　若 $V_n(F)$ 中两个向量组 $\{\alpha_1, \cdots, \alpha_n\}$ 与 $\{\beta_1, \cdots, \beta_n\}$ 等价,则 $L\{\alpha_1, \cdots, \alpha_n\} = L\{\beta_1, \cdots, \beta_n\}$.

定理 3.1.7　令 S 是 $V(F)$ 的一个向量集,以下各条等价:

1) S 线性无关且生成 V.

2) S 是一个极小生成集,如果 S 生成 V,且 S 的任意真子集不生成 V.

证明　1) $\Rightarrow$2). 由 1) 知 S 是一个生成集. 如果 S 的某个真子集 S' 也生成 V,那么 $S - S'$ 中的任意向量将会是 S' 中向量的一个线性组合,与 S 中向量都线性无关这一结论矛盾. 因此 2) 成立.

2) $\Rightarrow$1). 如果 S 是一个极小生成集,那么它一定线性无关. 因为如果 S 线性相关,那么任意向量 $s \in S$ 将会是 S 中其他向量的一个线性组合,因而 $S - \{s\}$ 将会是 S 的一个真生成子集,而这是不可能的. 所以 1) 成立.

定理 3.1.8　设 $\{\alpha_1, \cdots, \alpha_r\}$ 为 $V_n(F)$ 一组线性无关的向量,则 V 中必有 $n - r$ 个向量 $\alpha_{r+1}, \cdots, \alpha_n$,使得 $\{\alpha_1 \cdots \alpha_r, \alpha_{r+1} \cdots \alpha_n\}$ 作成 V 的一个基(也称为基的扩充定理).

例 3.1.2　n 维向量空间 $V_n(F)$ 的任意一个子空间都是某一含 n 个未知量的齐次线性方程组的解空间.

证明　设 $\{\alpha_1, \cdots, \alpha_n\}$ 为 $V_n(F)$ 的一个基,W 为 $V_n(F)$ 的任一 r 维子空间,其中 $0 \leqslant r \leqslant n$,令 $W = L(\beta_1, \cdots, \beta_r)$,于是有

$$\beta_i = \sum_{j=1}^{n} a_{ij}\alpha_j,\ i = 1, 2 \cdots r.$$

再令 $A = (a_{ij})_{r \times n}$,则 A 的秩为 r,这样以 A 为系数阵的齐次线性方程组 $AX = 0$ 的基础解系含有 $n - r$ 个线性无关的解向量 $\eta_k = (b_{k1}, \cdots, b_{kn})$,$k = 1, \cdots, n - r$,对此再以 $\eta_k (k = 1, \cdots, n - r)$ 为行构成一个 $(n - r) \times n$ 矩阵 B,则秩 $B = n - r$. 于是又以 B 为系数阵得到一齐次方程组 $BY = 0$,其解空间的维数为 r,由此可见,$BA' = 0$. 此式恰好说明了 A 的行空间等于齐次方程组 $BY = 0$ 的解空间.

然而,矩阵 A 的第 $i (i = 1, \cdots, r)$ 行是子空间 $W = L(\beta_1, \cdots, \beta_r)$ 的每一生成元关于基 $\{\alpha_1, \cdots, \alpha_n\}$ 的坐标,故 V 的子空间 $W = L(\beta_1, \cdots, \beta_r)$ 是由齐次线性方程组 $BY = 0$ 的解(向量)为坐标的全体向量构成的,即 W 为齐次线性方程组 $BY = 0$ 的解空间.

3.2　子空间的直和分解

3.2.1　坐标和过渡阵

定义 3.2.1　设 $\{\alpha_1, \cdots, \alpha_n\}$ 为 $V_n(F)$ 的一个基,$\forall \alpha \in V_n(F)$,有 $\alpha = k_1\alpha_1 + \cdots +$

$k_n\alpha_n$,则称这个有序数组 $k_1, \cdots, k_n$ 为 α 关于基 $\{\alpha_1, \cdots, \alpha_n\}$ 的**坐标**,记为$(k_1,\cdots,k_n)$.

定义 3.2.2 设 $V_n(F)$ 的两个基:$\{\alpha_1,\cdots,\alpha_n\}$,$\{\beta_1,\cdots,\beta_n\}$,且 $\beta_j = \sum_{i=1}^{n} a_{ij}\alpha_i$, $j = 1, 2 \cdots n$,则称 n 阶矩阵

$$T = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

为由基$\{\alpha_1,\cdots,\alpha_n\}$到基$\{\beta_1,\cdots,\beta_n\}$的**过渡矩阵**,简称**过渡阵**,记为$\{\beta_1,\cdots,\beta_n\} = \{\alpha_1,\cdots,\alpha_n\}T$.

定理 3.2.1(坐标变换公式) 设向量 ξ 关于 $V_n(F)$ 的两个基$\{\alpha_1, \cdots, \alpha_n\}$,$\{\beta_1, \cdots, \beta_n\}$ 的坐标分别为$(x_1, \cdots, x_n)'$ 与$(y_1, \cdots, y_n)'$,若$\{\beta_1, \cdots, \beta_n\} = \{\alpha_1, \cdots, \alpha_n\}T$,其中 T 为由基(1)$\{\alpha_1, \cdots, \alpha_n\}$到基(2)$\{\beta_1, \cdots, \beta_n\}$的过渡阵,则$(x_1, \cdots, x_n)' = (y_1, \cdots, y_n)'T$.

定理 3.2.2 $V_n(F)$ 中由基$\{\alpha_1, \cdots, \alpha_n\}$到基$\{\beta_1, \cdots, \beta_n\}$的过渡阵 T 是可逆的.

因此有以下三个重要结论:

(1) 由基$\{\beta_1,\cdots,\beta_n\}$到基$\{\alpha_1,\cdots,\alpha_n\}$的过渡阵为 T^{-1};

(2) 数域 F 上任一 n 阶可逆阵 A 都可作为 $V_n(F)$ 中一个基到另一个基的过渡阵;

(3) 设$\{\alpha_1,\cdots,\alpha_n\}$与$\{\beta_1,\cdots,\beta_n\}$分别为 $V_n(F)$ 的两组向量,$A \in M_n(F)$,并且

$$(\beta_1,\cdots,\beta_n) = (\alpha_1,\cdots,\alpha_n)A,$$

那么$\{\beta_1, \cdots, \beta_n\}$为 $V_n(F)$ 的一个基 $\Leftrightarrow\{\alpha_1,\cdots,\alpha_n\}$为 $V_n(F)$ 的一个基,且 A 可逆.

3.2.2 子空间的直和分解

定义 3.2.3 设 W,W' 分别是 $V(F)$ 的子空间,若

1) $W + W' = V$;

2) $W \cap W' = \{0\}$,

则称 W' 为 W 的一个余子空间,并称 V 是子空间 W 与 W' 的**直和**,记为 $V = W \oplus W'$.

一般地,设 W_1,W_2 为 $V(F)$ 的两个子空间,若 $W_1 + W_2$ 中每一向量 $\alpha = \alpha_1 + \alpha_2$, $\alpha_i \in W_i$, $i = 1, 2$,且这种表法唯一,则称 $W_1 + W_2$ 为 W_1 与 W_2 的直和,记为 $W_1 + W_2 = W_1 \oplus W_2$.

定理 3.2.3 设 V_1,V_2 为数域 F 上 n 维向量空间 $V_n(F)$ 的两个 $m(0 < m < n)$ 维子空间,则存在 V 的子空间 W,使 $V = V_1 \oplus W = V_2 \oplus W$.

证明 由题设 $\dim V_1 = \dim V_2 = m$ 且 $0 < m < \dim V = n$,知 V_1,V_2 为 V 的两个非平凡子空间,则存在 $\xi_1 \in V$,使 $\xi_1 \notin V_1$,$\xi_1 \notin V_2$.

设 $V_1 = L(\alpha_1, \cdots, \alpha_m)$，其中 $\alpha_1, \cdots, \alpha_m$ 为 V_1 的基，$V_2 = L(\beta_1, \cdots, \beta_m)$，其中 $\beta_1, \cdots, \beta_m$ 为 V_2 的基，于是 $\{\alpha_1, \cdots, \alpha_m, \xi_1\}$ 与 $\{\beta_1, \cdots, \beta_m, \xi_1\}$ 为 V 中两个线性无关的向量组. 令 $W_{11} = L(\alpha_1, \cdots, \alpha_m, \xi_1)$，$W_{21} = L(\beta_1, \cdots, \beta_m, \xi_1)$.

若 $m + 1 = n$，则令 $W = L(\xi_1)$，便有 $V = V_1 \oplus W = V_2 \oplus W$，即为所求.

若 $m + 1 < n$，则 W_{11}，W_{21} 还是 V 的两个非平凡子空间，于是存在 $\xi_2 \in V$，使 $\xi_2 \notin W_{11}$，$\xi_2 \notin W_{21}$，这样又有两个线性无关的向量组

$$\{\alpha_1, \cdots, \alpha_m, \xi_1, \xi_2\} \text{ 与 } \{\beta_1, \cdots, \beta_m, \xi_1, \xi_2\}$$

对此再令

$$W_{12} = L(\alpha_1, \cdots, \alpha_m, \xi_1, \xi_2),$$

$$W_{22} = L(\beta_1, \cdots, \beta_m, \xi_1, \xi_2),$$

它们均为 $m + 2$ 维子空间. 若 $m + 2 = n$，则 $L(\xi_1, \xi_2) = W$，即为所求. 若 $m + 2 < n$，则按上述思想，以此类推，由于 n 为有限数，故在 V 中必有 $\xi_1, \cdots, \xi_{n-m}$，使它们均不在 V_1，V_2 中，且又线性无关. 故令 $W = L(\xi_1, \cdots, \xi_{n-m})$ 为一个 $n - m$ 维子空间，使 $V = V_1 \oplus W = V_2 \oplus W$.

例 3.2.1 设 $A \in M_n(F)$，且 A 可逆，令 $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$. 证明：$n$ 元齐次方程组 $A_1 X = 0$ 与 $A_2 X = 0$ 的两个解空间的直和是 F^n.

证明 设 n 元齐次线性方程组 $AX = 0$，$A_1 X = 0$，$A_2 X = 0$ 的解空间分别是 W，W_1，W_2. 首先，W_1，W_2 皆为 F^n 的子空间是显然的，且 $W = W_1 \cap W_2$. 这是因为 $X \in W \Leftrightarrow AX = 0 \Leftrightarrow A_1 X = 0$ 且 $A_2 X = 0 \Leftrightarrow X \in W_1 \cap W_2$. 因 A 为可逆阵，以 A 为系数阵的 n 元齐次方程组的解空间 $W = \{0\}$，故有 $W_1 \cap W_2 = \{0\}$. 因此 $W_1 + W_2 = W_1 \oplus W_2$. 又由于 W_1，W_2 皆为 F^n 的子空间，则有 $W_1 \oplus W_2 \subseteq F^n$，又因

$$\dim W_1 = n - R(A_1), \dim W_2 = n - R(A_2) = n - (n - R(A_1)) = R(A_1),$$

有 $\dim W_1 + \dim W_2 = n$，即 $\dim(W_1 \oplus W_2) = \dim F^n$，故 $W_1 \oplus W_2 = F^n$，即 F^n 是 n 元齐次方程组 $A_1 X = 0$ 与 $A_2 X = 0$ 的解空间的直和.

3.3 向量空间的同构

3.3.1 向量空间的同构

定义 3.3.1 设 V 和 W 是数域 F 上的两个向量空间，若

1) f 是 V 到 W 的一个双射；

2) $\forall \xi, \eta \in V$，$f(\xi + \eta) = f(\xi) + f(\eta)$；

3) $\forall \xi \in V, \forall a \in F, f(a\xi) = af(\xi)$,

则称 f 为 V 到 W 的一个**同构映射**. 此时称 V 与 W **同构**,记为 $V \cong W$.

定理 3.3.1　设 V 和 W 是数域 F 上两个向量空间, f 是 V 到 W 的一个同构映射,那么

1) $f(0) = 0$;

2) $\forall \xi, \eta \in V, f(\xi + \eta) = f(\xi) + f(\eta)$;

3) $\forall \xi \in V, \forall a \in F, f(a\xi) = af(\xi)$;

4) $\forall \xi_i \in V, a_i \in F, f(a_1\xi_1 + a_2\xi_2 + \cdots + a_n\xi_n) = a_1 f(\xi_1) + a_2 f(\xi_2) + \cdots + a_n f(\xi_n)$;

5) $\xi_1, \xi_2, \cdots, \xi_n \in V$ 线性相关 $\Leftrightarrow f(\xi_1), f(\xi_2), \cdots, f(\xi_n)$ 线性相关;

6) f 的逆映射 f^{-1} 是 W 到 V 的同构映射.

定理 3.3.2　$V(F)$ 与 $W(F)$ 同构 $\Leftrightarrow \dim V = \dim W$.

任意 n 维空间 V 都与 F^n 同构,所以 $F, F^2, \cdots, F^n, \cdots$ 代表了所有向量空间.

例 3.3.1　设 a,b 为两个复数,令

$$V_a = \{f(x) \mid f(x) \in F[x], f(a) = 0\}, V_b = \{g(x) \mid g(x) \in F[x], g(b) = 0\},$$

为 $F[x]$ 的两个子空间,试证: V_a 与 V_b 同构.

证明　由题设和 V_a 与 V_b 中多项式的性质来建立它们之间的一个同构映射. 规定

$$\varphi: f(x) = (x-a)h(x) \to g(x) = (x-b)h(x).$$

其中 $\forall f(x) \in V_a$. 显然 φ 是 V_a 到 V_b 的一个映射,并且是一个双射.

对于 $\forall f_1(x), f_2(x) \in V_a$,且 $f_1(x) \neq f_2(x)$,则由 $f_1(x) = (x-a)h_1(x), f_2(x) = (x-a)h_2(x)$ 得 $h_1(x) \neq h_2(x)$,因此 $\varphi(f_1(x)) = (x-b)h_1(x) \neq (x-b)h_2(x) = \varphi(f_2(x))$,故 φ 为单射;

$\forall g(x) \in V_b$,由于 $g(x) = (x-b)h(x)$,则有 $f(x) = (x-a)h(x) \in V_a$,使 $\varphi(f(x)) = (x-b)h(x) = g(x)$,故 φ 为满射,进而 φ 为 V_a 到 V_b 的一个双射.

同时, $\forall f_1(x), f_2(x) \in V_a, \forall k, l \in F$,由于

$$f_1(x) = (x-a)h_1(x), f_2(x) = (x-a)h_2(x),$$

$$\varphi(f_1(x)) = (x-b)h_1(x), \varphi(f_2(x)) = (x-b)h_2(x),$$

得

$$\begin{aligned}\varphi(kf_1(x) + lf_2(x)) &= \varphi(k(x-a)h_1(x) + l(x-a)h_2(x)) \\ &= \varphi(x-a)(kh_1(x) + lh_2(x)) \\ &= (x-b)(kh_1(x) + lh_1(x)) \\ &= k\varphi(f(x)) + l\varphi(f_2(x)).\end{aligned}$$

综上所述,φ 确为 V_a 与 V_b 之间的一个同构映射,因而 V_a 与 V_b 同构.

3.3.2 商空间

定义 3.3.2 令 S 是 V 的一个子空间,且令 $\equiv_S$ 为 V 上由 $u \equiv_S v \Leftrightarrow u - v \in S$ 定义的二元关系,易知 $\equiv_S$ 是一个等价关系. 当 $u \equiv_S v$ 时,我们就说 u 和 v 是**模 S 同余的**,常用术语 mod 代替 modulo,并且 $u \equiv_S v$ 记为 $u \equiv v \pmod S$,简记为 $u \equiv v$.

为了研究等价类的外部特征,注意到

$$\begin{aligned}[v] &= \{u \in V \mid u \equiv v\} = \{u \in V \mid u - v \in S\} \\ &= \{u \in V \mid u = v + s \text{ 对于 } \forall s \in S\} \\ &= \{v + s \mid s \in S\} = v + S.\end{aligned}$$

在 V 中,集合 $f = v + S = \{v + s \mid s \in S\}$ 叫做 S 的一个**陪集**.

定义 3.3.3 在 V 中,模 S 同余的等价类是 S 的陪集 $v + S$. 全体陪集所成的集合表示为 $\dfrac{V}{S} = \{v + S \mid v \in V\}$,读作"$V$ mod S",叫做 **V 模 S 的商空间**.

商空间的性质如下:

1) $u + S = v + S \Leftrightarrow v \in u + S \Leftrightarrow u \in v + S$;$u, v \in V \Rightarrow u + S = v + S$ 或 $(u + S) \cap (v + S) = \varnothing$.

因此,对于不同的向量 v,陪集可以表示成形式 $v + S$.

2) 当陪集 f 记作 $v + S$ 时,向量 v 叫做 f 的**陪集代表**. 显然,陪集中任意向量都能作为陪集代表.

同样注意到 $u \equiv v \Rightarrow u - v \in S \Rightarrow r(u - v) \in S$,对于 $\forall r \in F \Rightarrow ru - rv \in S$,对于 $\forall r \in F \Rightarrow ru \equiv rv$. 所以 $u \equiv v \Rightarrow ru \equiv rv$,对于 $\forall r \in F$.

3) 另外,$u_1 \equiv v_1, u_2 \equiv v_2 \Rightarrow u_1 - v_1 \in S, u_2 - v_2 \in S \Rightarrow (u_1 - v_1) + (u_2 - v_2) \in S$

$$\Rightarrow (u_1 + u_2) - (v_1 + v_2) \in S \Rightarrow (u_1 + u_2) \equiv (v_1 + v_2).$$

所以 $u_1 \equiv v_1, u_2 \equiv v_2 \Rightarrow (u_1 + u_2) \equiv (v_1 + v_2)$.

由性质 2) 和 3) 可知模 S 同余保持 V 上的向量空间运算. 它们与陪集代表的选取无关,即 $u_1 + S = u_2 + S, v_1 + S = v_2 + S \Rightarrow (u_1 + v_1) + S = (u_2 + v_2) + S$;$u_1 + S = u_2 + S \Rightarrow r(u_1 + S) = r(u_2 + S)$.

定义 3.3.4 令 S 是 V 的一子空间,对于 $\forall v \in S$,由 $\pi_S(v) = v + S$,定义映射 π_S: $V \to V/S$. 这一映射叫做从 V 到 S 上的**典范射影**,或**自然射影**.

容易看出它是线性的,

$$\pi(ru + sv) = (ru + sv) + S = r(u + S) + s(v + S) = r\pi(u) + s\pi(v).$$

典范射影是满射的,因为对于任意陪集 $v + S$,$v + S = \pi(v)$. 为了确定 π 的核,有

$$v \in \operatorname{Ker}(\pi) \Leftrightarrow \pi(v) = 0 \Leftrightarrow v + S = S \Leftrightarrow v \in S,$$

所以 $\operatorname{Ker}(\pi) = S$.

图 3-1 说明存在从商空间 V/S 到 W 的映射 τ',有性质

4) $\tau' \cdot \pi_S = \tau$,即 $\tau(v) = (\tau' \cdot \pi_S)(v) = \tau'(v+S)$.

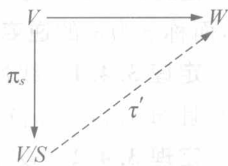


图 3-1

因此,由 $\tau'(v+S) = \tau(v)$,我们定义一个从 V/S 到 W 的函数 τ',这个函数是唯一确定的,当且仅当 $v+S = u+S \Rightarrow \tau'(v+S) = \tau'(u+S)$,或等价地,$v+S = u+S \Rightarrow \tau(v) = \tau(u)$. 而这又等价于 $v-u \in S \Rightarrow \tau(v-u) = 0$. 或用 x 代替 $v-u$,则 $x \in S \Rightarrow \tau(x) = 0$. 因此,$\tau'$ 是唯一确定的,当且仅当 $S \subset \operatorname{Ker}(\tau)$. 我们假设 $S \subset \operatorname{Ker}(\tau)$,因此 τ' 是唯一确定的. 那么 τ': $V/S \to W$ 是一线性变换,有象

$$\operatorname{Im}(\tau') = \{\tau'(v+S) \mid v+S \in V/S\} = \{\tau(v) \mid v \in V\} = \operatorname{Im}(\tau)$$

和核

$$\operatorname{Ker}(\tau') = \{v+S \mid \tau'(v+S) = 0\} = \{v+S \mid \tau(v) = 0\} = \{v+S \mid v \in \operatorname{Ker}(\tau)\}.$$

同样,由于存在唯一的映射 τ': $V/S \to W$ 有性质 $\tau' \cdot \pi_S = \tau$,所以 τ' 是唯一的.

定理 3.3.3 令 S 为 V 的一子空间,那么在 V 中,S 的任意补子空间都同构于商空间 $\dfrac{V}{S}$.

定理 3.3.4(第一同构定理) 令 τ: $V \to W$ 是一个线性变换,那么由 $\varphi(v+\operatorname{Ker}(\tau)) = \tau(v)$ 定义的线性变换 φ: $V/\operatorname{Ker}(\tau) \to W$ 是单射,从而 $V/\operatorname{Ker}(\tau) \approx \operatorname{Im}(\tau)$.

3.4 线性变换

3.4.1 线性变换的定义及运算

定义 3.4.1 数域 F 上向量空间 V 的一个变换 σ 称为**线性变换**,如果对 $\forall \alpha, \beta \in V$, $\forall k \in F$,都有 ① $\sigma(\alpha+\beta) = \sigma(\alpha) + \sigma(\beta)$;② $\sigma(k\alpha) = k\sigma(\alpha)$,即 σ 满足齐次可加性.

设 $L(V)$ 为数域 F 上向量空间 V 的所有线性变换构成的集合.

1) 加法:$\forall \sigma, \tau \in L(V)$, $\forall \alpha \in V$,令 $(\sigma+\tau)(\alpha) = \sigma(\alpha) + \tau(\alpha)$,易知 $\sigma+\tau \in L(V)$,称 $\sigma+\tau$ 为 σ 与 τ 的**和**;

2) 数量乘法:$\forall \sigma \in L(V)$, $\forall k \in F$, $\forall \alpha \in V$,令 $(k\sigma)(\alpha) = k(\sigma(\alpha))$,易知 $k\sigma \in L(V)$,称 $k\sigma$ 为数 k 与线性变换 σ 的**数积**;

$L(V)$ 对上述两种运算构成 F 上一个向量空间.

3) 乘法:$\forall \sigma, \tau \in L(V)$, $\forall \alpha \in V$,令 $(\sigma\tau)(\alpha) = \sigma(\tau(\alpha))$,易知 $\sigma\tau \in L(V)$,称 $\sigma\tau$

为σ与τ的积．线性变换乘法运算满足结合律、乘法对加法的分配律，但不满足交换律．

定义 3.4.2 设$\sigma \in L(V)$，若$\tau \in L(V)$，使$\sigma\tau = \sigma\tau = \iota$（恒等变换），则称$\sigma$是可逆的，而称$\tau$为$\sigma$的**逆变换**，记$\tau = \sigma^{-1}$，易知$\sigma^{-1} \in L(V)$．

定理 3.4.1 线性变换将线性相关的向量组变成线性相关的向量组（注意，反之不然），且$\sigma L(\alpha_1, \cdots, \alpha_n) = L(\sigma(\alpha_1), \cdots, \sigma(\alpha_n))$．

定理 3.4.2 设$\{\alpha_1, \cdots, \alpha_n\}$为$V_n(F)$的一个基，则对任意$V$中任意给定$n$个向量$\beta_1, \cdots, \beta_n$，必确定唯一$\sigma \in L(V)$，使$\sigma(\alpha_i) = \beta_i, i = 1, \cdots, n$．

定理 3.4.3 数域F上n维向量空间的线性变换集合构成的向量空间$L(V_n)$与$M_n(F)$同构，即$L(V_n) \cong M_n(F)$．

定理 3.4.4 令A是F上一$m \times n$矩阵．

1）τ_A是单射的，当且仅当$rk(A) = n$；

2）τ_A是满射的，当且仅当$rk(A) = m$．

3.4.2 线性变换的象与核

定义 3.4.3 设$\sigma \in L(V)$，令$\mathrm{Im}(\sigma) = \{\sigma(\xi) \mid \xi \in V\} = \sigma(V)$，$\mathrm{Ker}(\sigma) = \{\xi \in V \mid \sigma(\xi) = 0\}$，分别称为$\sigma$的**象与核**，易知$\sigma$的象与核都是$V$的子空间．

定理 3.4.5 设$\alpha_1, \cdots, \alpha_n$为$V_n(F)$的基，$\sigma \in L(V)$，$A$是$\sigma$关于基$\{\alpha_1, \cdots, \alpha_n\}$的矩阵，则

1）$\mathrm{Im}(\sigma) = L(\sigma(\alpha_1), \cdots, \sigma(\alpha_n))$；

2）$\dim\mathrm{Im}(\sigma) = R(A)$；

3）$\dim\mathrm{Im}(\sigma) + \dim\mathrm{Ker}(\sigma) = n$；

通常称$\dim\mathrm{Im}(\sigma)$为σ的**秩**，称$\dim\mathrm{Ker}(\sigma)$为σ的**零度**；

4）σ为单射$\Leftrightarrow \mathrm{Ker}(\sigma) = \{0\}$；$\sigma$为满射$\Leftrightarrow \mathrm{Im}(\sigma) = V$；$\sigma$为单射$\Leftrightarrow \sigma$为满射．

定理 3.4.6 设$\sigma \in L(V)$，下列诸命题彼此等价：

1）有$\tau \in L(V)$使$\sigma\tau = \tau\sigma = \iota$（$\iota$为恒等变换），称$\sigma$为可逆的；

2）σ关于V的任意一基的矩阵A必可逆；

3）设$\alpha_1, \cdots, \alpha_n$为$V$的一个基，则$\sigma(\alpha_1), \cdots, \sigma(\alpha_n)$仍为一个基；

4）σ的秩$\dim\mathrm{Im}(\sigma) = n$；$\sigma$的零度$\dim\mathrm{Ker}(\alpha) = 0$．

5）$\mathrm{Im}(\sigma) = V$；$\mathrm{Ker}(\sigma) = \{0\}$；

6）对$\forall \alpha, \beta \in V_n(F)$，当$\alpha \neq \beta$时，必有$\sigma(\alpha) \neq \sigma(\beta)$；

7）存在一个常数项不为零的多项式$f(x)$，使$f(\sigma) = 0$；

8）σ的特征值（根）均不为零；

9）若$V = V_1 \oplus V_2$，则$V = \sigma(V_1) \oplus \sigma(V_2)$．

例 3.4.1　设 V_1, V_2 是 $V_n(F)$ 的两个子空间，且 $\dim V_1 + \dim V_2 = \dim V = n$. 试证：必存在一个线性变换 σ，使 $\mathrm{Im}(\sigma) = V_2$，$\mathrm{Ker}(\sigma) = V_1$.

证明　令 $\dim V_1 = r$，则 $\dim V_2 = n - r$，设 $V_1 = L(\alpha_1, \cdots, \alpha_r)$，$V_2 = L(\beta_{r+1}, \cdots, \beta_n)$，其中 $\alpha_1, \cdots, \alpha_r$ 线性无关，$\beta_{r+1}, \cdots, \beta_n$ 也线性无关. 由 $\alpha_1, \cdots, \alpha_r$ 扩充为 V 的一个基 $\{\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n\}$，对 V 中 n 个向量 $0, \cdots, 0, \beta_{r+1}, \cdots, \beta_n$，必有一个线性变换 σ，使 $\sigma(\alpha_i) = 0, i = 1, \cdots, r, \sigma(\alpha_i) = \beta_i: i = r+1, \cdots, n$，

下证 $V_2 = \mathrm{Im}(\sigma)$，$V_1 = \mathrm{Ker}(\sigma)$，事实上

$$\begin{aligned}\mathrm{Im}(\sigma) &= \sigma(L(\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n)) \\ &= L(\sigma(\alpha_1), \cdots, \sigma(\alpha_r), \sigma(\alpha_{r+1}), \cdots, \sigma(\alpha_n)) \\ &= L(\beta_{r+1}, \cdots, \beta_n) = V_2.\end{aligned}$$

又 $\forall \alpha \in V_1 = L(\alpha_1, \cdots, \alpha_r)$，$\alpha = \sum_{i=1}^{r} k_i \alpha_i$，于是 $\sigma(\alpha) = \sigma(\sum_{i=1}^{r} k_i \alpha_i) = \sum_{i=1}^{r} k_i \sigma(\alpha_i) = 0$，即得 $\sigma \in \mathrm{Ker}(\sigma)$，即 $V_1 \subseteq \mathrm{Ker}(\sigma)$. 又 $\dim\mathrm{Ker}(\sigma) = n - \dim\mathrm{Im}(\sigma) = n - (n - r) = r = \dim V_1$，因此 $\mathrm{Ker}(\sigma) = V_1$.

定理 3.4.7　令 $\tau \in \Lambda(V, W)$. 我们有以下定理：

1)（秩加零度定理）　$\dim(\mathrm{Ker}(\tau)) + \dim(\mathrm{Im}(\tau)) = \dim(V)$ 或 $\mathrm{rk}(\tau) + \mathrm{null}(\tau) = \dim(V)$.

2) 若 S 是 V 的一个子空间，那么 S 的全体补子空间都是同构的.

3) 若 S 是 V 的一个子空间，那么 $\dim(S) + \dim(S^c) = \dim(V)$.

推论 3.4.8　令 $\tau \in \Lambda(V, W)$，且 $\dim(V) = \dim(W) < \infty$，则 τ 是单射 $\Leftrightarrow$ 是满射.

3.4.3　线性变换的矩阵

定义 3.4.4　设 $\{\alpha_1, \cdots, \alpha_n\}$ 为 $V_n(F)$ 的一个基，$\sigma \in L(V)$，若 $\sigma(\alpha_j) = a_{1j}\alpha_1 + a_{2j}\alpha_2 + \cdots + a_{nj}\alpha_n, j = 1, 2, \cdots, n$，则以每一基向量 $\alpha_j (j = 1, 2, \cdots, n)$ 在 σ 之下的象 $\sigma(\alpha_j)$ 关于基 $\{\alpha_1, \cdots, \alpha_n\}$ 的坐标作为列构成的一个 n 阶方阵

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

称为 σ 关于基 $\{\alpha_1, \cdots, \alpha_n\}$ 的矩阵. 通常有 $(\sigma((\alpha_1), \cdots, \sigma(\alpha_n)) = (\alpha_1, \cdots, \alpha_n)A$.

定理 3.4.9　设 $\sigma \in L(V)$，$\{\alpha_1, \cdots, \alpha_n\}$ 为 V 的一个基，A 为 σ 关于基 $\{\alpha_1, \cdots, \alpha_n\}$ 的矩阵. 若向量 ξ 关于基 $\{\alpha_1, \cdots, \alpha_n\}$ 的坐标是 $(x_1, \cdots, x_n)$，则 $\sigma(\xi)$ 关于基 $(\alpha_1, \cdots, \alpha_n)$ 的坐标是 $(y_1, \cdots, y_n)$ 可按公式 $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ 计算.

例 3.4.2 在 V_2 中取从原点出发的两个彼此正交的单位向量 ε_1，ε_2 作为 V_2 的基，$\sigma \in L(V_2)$，使 V_2 的每一个向量旋转 θ 角的变换，求 σ 关于基 $\{\varepsilon_1, \varepsilon_2\}$ 的矩阵.

解 因为 $\sigma(\varepsilon_1) = \cos\theta\varepsilon_1 + \sin\theta\varepsilon_2$；$\sigma(\varepsilon_2) = -\sin\theta\varepsilon_1 + \cos\theta\varepsilon_2$，

所以 σ 关于基 $\{\varepsilon_1, \varepsilon_2\}$ 的矩阵是 $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$；

设 ξ 关于基 $\{\varepsilon_1, \varepsilon_2\}$ 的坐标是 $\{x_1, x_2\}$，而 $\sigma(\xi)$ 关于基 $\{\varepsilon_1, \varepsilon_2\}$ 的坐标是 $\{y_1, y_2\}$，则

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

定理 3.4.10 设 V 是数域 F 上一个 n 维向量空间，$\{\alpha_1, \cdots, \alpha_n\}$ 为 V 的一个基，那么对于 V 中任意向量 $\beta_1, \cdots, \beta_n$，恰有 $\sigma \in L(V)$，设 $\sigma(\alpha_i) = \beta_i$，$i = 1, 2, \cdots, n$.

定理 3.4.11 设 V 是数域 F 上一个 n 维向量空间，$\{\alpha_1, \cdots, \alpha_n\}$ 为 V 的一个基，对于 V 的每一个 $\sigma \in L(V)$，令 σ 关于基 $\{\alpha_1, \cdots, \alpha_n\}$ 的矩阵 A 与之对应，这样就得到 $L(V_i)$ 到 $M_n(F)$ 的一个双射，并且 $\sigma, \tau \in L(V)$，而 $\sigma \to A$，$\tau \to B$，那么 $\sigma + \tau \to A + B$，$a\sigma \to aA$，$\sigma\tau \to AB$.

如果 σ 可逆，那么 A 可逆，并且 $\sigma^{-1} \to A^{-1}$.

定义 3.4.5 设 $A, B \in M_n(F)$，如果存在 $T \in M_n(F)$，使 $B = T^{-1}AT$，称 B 与 A **相似**.

矩阵相似是一个等价关系.

定理 3.4.12 设 $\sigma \in L(V)$，σ 关于 V 的两个基 $\{\alpha_1, \cdots, \alpha_n\}$ 与 $\{\beta_1, \cdots, \beta_n\}$ 的矩阵分别为 A 与 B，并且 $(\beta_1, \cdots, \beta_n) = (\alpha_1, \cdots, \alpha_n)T$，其中 T 为由基 $\{\alpha_1, \cdots, \alpha_n\}$ 到基 $\{\beta_1, \cdots, \beta_n\}$ 的过渡阵，则 $B = T^{-1}AT$，即 A 与 B 相似.

3.5 线性变换的对角化

3.5.1 不变子空间与约化对

令 τ 是 V 上一线性算子，如果 S 是 V 的一个子空间，那么对于一取定的 $s \in S$，向量 $\tau(s)$ 也在 S 内这一条件并不一定要满足，这就有

定义 3.5.1 设 $\sigma \in L(V)$，W 为 V 的子空间，若 $\forall \alpha \in W$，都有 $\sigma(\alpha) \in W$，则称 W 是 σ 的**不变子空间**，称 W 为 σ -子空间.

易知 $\{0\}$，V，$\mathrm{Im}(\sigma)$，$\mathrm{Ker}(\sigma)$ 都是每一线性变换 σ 的不变子空间，σ 的属于任一特征值 λ 的特征子空间 V_λ 也是 σ -子空间.

定义 3.5.2 令 τ 是 V 上一线性算子，如果 $V = S \oplus T$，且若 S 和 T 在 τ 之下都是

不变的，我们就说 (S,T) **约化** τ，即 (S,T) 约化 τ，如果限制 $\tau|_S$ 和 $\tau|_T$ 分别是 S 和 T 上的线性算子.

定义 3.5.3　令 ρ 是 V 上一线性算子，记作 $\rho=\sigma\oplus\tau$，且称 ρ 是 σ 和 τ 的**直和**，如果存在 V 的子空间 S 和 T，使得 (S,T) 约化 ρ，且 $\Sigma=\rho|_S,\tau=\rho|_T$.

线性算子的直和这一概念对线性算子结构的研究起着重要作用. 本书后面会详细讨论到：如果 $\rho=\sigma\oplus\tau$，那么我们可以把 ρ 分解成更为简单的线性算子 σ 和 τ.

例 3.5.1　设 $V=V_1\oplus V_2,\sigma,\tau\in L(V),\forall\alpha=\alpha_1+\alpha_2\in V(\alpha_i\in V_i,i=1,2)$ 都有 $\sigma(\alpha)=o_1$，求证 V_1 与 V_2 都是 τ 的不变子空间的充要条件是 $\sigma\tau=\tau\sigma$.

证明　*必要性*　$\forall\alpha=\alpha_1+\alpha_2\in V(\alpha_i\in V_i,i=1,2)$，则 $\tau(\alpha)=\tau(\alpha_1)+\tau(\alpha_2)\in V=V_1\oplus V_2$，进而 $\sigma(\tau(\alpha))=\tau(\alpha_1),\tau(\sigma(\alpha))=\tau(\alpha_1)$，故 $\sigma(\tau(\alpha))=\tau(\sigma(\alpha))$，即 $\sigma\tau=\sigma\tau$.

充分性　$\forall\alpha\in V_1$，则 $\sigma(\alpha)=\alpha,\sigma(\tau(\alpha))=\tau(\sigma(\alpha))=\tau(\alpha)$，令 $\tau(\alpha)=\beta_1+\beta_2\in V,\beta_i\in V_i,i=1,2$，于是有 $\tau(\alpha)=\sigma(\tau(\alpha))=\sigma(\beta_1+\beta_2)=\beta_1$，即 $\tau(V_1)\subseteq V_1$，故 V_1 是 τ 的不变子空间.

同理 $\forall\alpha\in V_2,\sigma(\alpha)=0$，则 $\sigma(\tau(\alpha))=\tau(\sigma(\alpha))=\tau(0)=0$，令 $\tau(\alpha)=\beta_1+\beta_2\in V,\beta_i\in V_i,i=1,2$，于是有 $\sigma(\tau(\alpha))=\sigma(\beta_1+\beta_2)=\beta_1$，由此得 $\beta_1=0$，因此 $\tau(\alpha)=\beta_2\in V_1$，即 $\tau(V_2)\subseteq V_2$，故 V_2 也是 τ 的不变子空间.

定理 3.5.1（Fitting 定理）　设 σ 为数域 F 上的 n 维向量空间 V 的线性变换，令 $V_1=\{\alpha\mid\alpha\in V$，有某一自然数 m，使 $\sigma^m(\alpha)=0\}$，$V_2=\bigcap_{i=1}^{\infty}\sigma^i(V)$. 求证：

1) V_1,V_2 都是 σ-子空间；2) $\sigma|_{V_1}$ 是幂零变换，$\sigma|_{V_2}$ 是可逆的；3) $V=V_1\oplus V_2$.

证明　1) 由 $V\supseteq\sigma(V)$，有 $\sigma(V)\supseteq\sigma^2(V)$，因此 $V\supseteq\sigma(V)\supseteq\sigma^2(V)\supseteq\cdots$，故必有某一自然数 s 使 $\sigma^s(V)=\sigma^{s+1}(V)=\cdots$，则 $V_2=\sigma^s(V)=\sigma^{s+1}(V)=\cdots$，故有自然数 t，使

$$\mathrm{Ker}(\sigma^t)\subseteq\mathrm{Ker}(\sigma^{t+1})\subseteq\cdots,$$

所以 $V_1=\mathrm{Ker}(\sigma^t)=\mathrm{Ker}(\sigma^{t+1})=\cdots$. 因此 V_1,V_2 皆为 σ-子空间.

2) 由 $\sigma|_{V_2}(V_2)=\sigma(V_2)=V_2$，显然，$\sigma|_{V_2}$ 为可逆的线性变换. 又对于 $\forall\alpha\in V_1$，有 $(\sigma|_{V_1})^t(\alpha)=\sigma^t(\alpha)=0$，即 $\sigma|_{V_1}$ 为幂等变换.

3) $\forall\xi\in V_1\cap V_2$，由 $\xi\in V_1$ 且 $\xi\in V_2$，有 $\sigma^m(\xi)=0$，又 $\sigma|_{V_2}$ 为可逆的，故 $\xi=0$，即 $V_1\cap V_2=\{0\}$. 又 $\sigma^s(V)=\sigma^{2s}(V)$，因此对于 $\forall\alpha\in V$，必有 $\beta\in V$，使得 $\sigma^s(\alpha)=\sigma^s(\beta)$，所以 $\sigma^s(\alpha-\sigma^s(\beta))=0$，于是 $\alpha-\sigma^s(\beta)\in V_1$，而 $\sigma^s(\beta)\in V_2$，故 $\alpha=(\alpha-\sigma^s(\beta))+\sigma^s(\beta)\in V_1+V_2$，即 $V\subseteq V_1+V_2$.

反之，显然有 $V\supseteq V_1+V_2$，进而 $V=V_1+V_2$.

例 3.5.2　设 σ 为数域 F 上 n 维向量空间 V 的线性变换，$\mathrm{Im}(\sigma)=\mathrm{Ker}(\sigma)$，试证：选择 V 的一个基，使 σ 关于这基的矩阵为 Jordan 标准形，并把标准形写出来.

证明　由题设 $\mathrm{Im}(\sigma)=\mathrm{Ker}(\sigma)$ 及维数公式 $\dim\mathrm{Im}(\sigma)+\dim\mathrm{Ker}(\sigma)=\dim V=n$，

知 V 的维数 n 为偶数，令 $n=2r$，有 $\dim\mathrm{Im}(\sigma)=r$，于是设 $\mathrm{Im}(\sigma)=L(\beta_1,\cdots,\beta_r)$，其中 $\beta_1,\cdots,\beta_r$ 线性无关，则存在 $\alpha_1,\cdots,\alpha_r\in V$，使得 $\sigma(\alpha_i)=\beta_i, i=1,\cdots,n$. 因为 $\mathrm{Im}(\sigma)=\mathrm{Ker}(\sigma)$，故有 $\sigma(\beta_i)=0, i=1,\cdots,n$. 设 $a_1\alpha_1+\cdots+a_r\alpha_r+b_1\beta_1+\cdots+b_r\beta_r=0$，对此式两边均用 σ 作用得 $a_1\sigma(\alpha_1)+\cdots+a_r\sigma(\alpha_r)=0$，即 $a_1\beta_1+\cdots+a_r\beta_r=0$，因 $\beta_1,\cdots,\beta_r$ 线性无关，有 $a_1=\cdots=a_r=0$，将此代入(1)式又得 $b_1\beta_1+\cdots+b_r\beta_r=0$，进而 $b_1=\cdots=b_r=0$，故 $2r$ 个向量 $\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_r$ 线性无关，因此构成 V 的一个基. 但是 σ 关于基 $\{\alpha_1,\cdots,\alpha_r,\beta_1,\cdots,\beta_r\}$ 的矩阵是 $\begin{bmatrix}0_r & 0_r\\ I_r & 0_r\end{bmatrix}$.

再选择 V 的一个基 $\{\alpha_1,\beta_1,\cdots,\alpha_r,\beta_r\}$，使 σ 关于该基的矩阵是 $J=\begin{bmatrix}J_1 & & \\ & \ddots & \\ & & J_r\end{bmatrix}$，其中 $J_i=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix}$, $i=1,\cdots,r$，即为所求.

3.5.2 特征根、特征向量、特征多项式

定义 3.5.4 线性变换 σ 的特征多项式 $f_\sigma(\lambda)$ 指的是 σ 关于某一基的矩阵 A 的**特征多项式** $f_A(\lambda)=|\lambda I-A|$.

定理 3.5.2 相似矩阵具有相同的特征多项式.

定理 3.5.3 设 $\sigma\in L(V)$，$f(\lambda)$ 是 σ 的特征多项式，则 $f(\sigma)=0$.

相平行地，关于 $A\in M_n(F)$，若 $f(\lambda)=|\lambda I-A|$ 是 A 的特征多项式，则

$$f(A)=A^n-Tr(A)A^{n-1}+\cdots+(-1)^n|A|I=0$$

此结论称为著名的 **Hamilton-Cayley 定理**.

定理 3.5.4 线性变换的属于不同特征值的特征向量是线性无关的. 因此，设 $\lambda_1\cdots\lambda_k$ 为线性变换 σ 的不同的特征值，而 $\alpha_{i1},\alpha_{i2},\cdots,\alpha_{in}$ 是属于特征值 λ_i 的线性无关的特征向量，$i=1,2,\cdots,k$，则向量组 $\alpha_{11},\cdots,\alpha_{1r},\cdots,\alpha_{k1},\cdots\alpha_{kr}$ 也线性无关.

定理 3.5.5 设线性变换 σ 的特征多项式 $f(\lambda)$ 可分解成一次因式方幂的乘积

$$f(\lambda)=(\lambda-\lambda_1)^{r_1}(\lambda-\lambda_2)^{r_2}\cdots(\lambda-\lambda_k)^{r_k}.$$

则 V 可分解成不变子空间的直和 $V=V_1\oplus V_2\oplus\cdots\oplus V_k$，其中

$$V_i=\{\xi\mid(\sigma-\lambda_i\varepsilon)^{r_i}\xi=0,\xi\in V\}, i=1,2,\cdots,k.$$

称 V_i 为 σ 的属于特征值 λ 的**根子空间**.

定理 3.5.6 设 σ 是复数域 **C** 上 n 维向量空间 V 的线性变换，则 V 中必有一个基，使 σ 关于这个基的矩阵是 Jordan 形阵，并且这个 Jordan 形阵除去其中 Jordan 块的排列

次序外是被 σ 唯一决定的.

定义 3.5.5　设 $\sigma \in L(V)$,$\lambda \in F$,若存在一非零向量 $\xi \in V$,使得 $\sigma(\xi) = \lambda\xi$,则称 λ 为 σ 的一个**特征根**,而 ξ 称为 σ 的属于特征值 λ 的**特征向量**.

定义 3.5.6　σ 的属于特征值 λ 的全部特征向量再添上零向量所成的集合,是 V 的一个子空间,称为 σ 的属于特征值 λ 的**特征子空间**,记为

$$V_\lambda = \{\alpha \mid \sigma(\alpha) = \lambda\alpha, \alpha \in V\}.$$

定义 3.5.7　线性变换 σ 可对角化：如果 σ 关于 V 的一个基 $\{\alpha_1, \cdots, \alpha_n\}$ 的矩阵是对角形 $\mathrm{diag}\{\lambda_1, \cdots, \lambda_n\}$.

n 阶矩阵 A 可对角化指的是存在 n 阶可逆阵 T,使 $T^{-1}AT$ 为对角形阵.

定理 3.5.7　设 σ 为数域 F 上 n 维向量空间 V 的线性变换,σ 可对角化与下列四个命题彼此等价：

1) $V = V_1 \oplus V_2 \oplus \cdots \oplus V_n$,其中 $V_i(i = 1, \cdots, n)$ 是 V 的一维 σ-子空间；

2) V 中有一个由 σ 的特征向量构成的基；

3) σ 的所有特征子空间的维数之和等于 n；

4) σ 的特征多项式的根都在 F 内,并且对 σ 的特征多项式的每一根 λ_i,特征子空间 V_{λ_i} 的维数等于 λ_i 的重数.

定理 3.5.8　设 V_1, V_2 为 $V_n(F)$ 的两个子空间,且 $V = V_1 \oplus V_2$,令 $\sigma: \xi \mapsto \xi_1$,其中 $\forall \xi = \xi_1 + \xi_2 \in V, \xi_1, \xi_2 \in V$,称 σ 为 V 的关于直和分解 $V = V_1 \oplus V_2$ 在 V_1 上的射影,简称 σ 是 V 在 V_1 上的射影. 证明

1) σ 是 V 的一个线性变换,即 $\sigma \in L(V_1)$；

2) $\mathrm{Im}(\sigma) = V_1, \mathrm{Ker}(\sigma) = V_2$,且 σ 在 V_1 上的限制 $\sigma|_{V_1}$ 是 V_1 的单位变换,即 $\sigma|_{V_1} = \varepsilon$;

3) $\sigma^2 = \sigma$,即 σ 为幂等变换；

4) $(\varepsilon - \sigma)$ 是 V 的关于直和分解 $V = \mathrm{Im}(\sigma) \oplus \mathrm{Ker}(\sigma)$ 到 $\mathrm{Ker}(\sigma)$ 的射影；

5) $V = \mathrm{Ker}(\sigma) \oplus \mathrm{Ker}(\varepsilon - \sigma) = \mathrm{Im}(\varepsilon - \sigma) \oplus \mathrm{Im}(\sigma)$；

6) $\mathrm{Im}(\varepsilon - \sigma), \mathrm{Ker}(\varepsilon - \sigma)$ 皆为 σ-子空间；

7) $\varepsilon + \sigma$ 为可逆线性变换；

8) σ 可对角化.

证明　1) $\forall \xi, \eta \in V, a, b \in F$,则有 $\xi = \xi_1 + \xi_2, \eta = \eta_1 + \eta_2$,其中 $\xi_i \in V_i, H_i \in V_i$, $i = 1, 2$. 于是 $\sigma(a\xi + b\eta) = \sigma(a\xi_1 + a\xi_2 + b\eta_1 + b\eta_2) = a\xi_1 + b\eta_1 = a\sigma(\xi) + b\sigma(\eta)$,故 σ 为线性变换.

2) $\forall \alpha \in \mathrm{Im}(\sigma)$,则有 $\beta \in V$,使 $\sigma(\beta) = \alpha$,然而 $\beta = \beta_1 + \beta_2 \in V, \beta_i \in V_i, i = 1, 2$,因此 $\alpha = \sigma(\beta) = \beta_i \in V_i$,即得 $\mathrm{Im}(\sigma) \subseteq V_1$;反之,$\forall \alpha_1 \in V_1, \alpha_1 = \alpha + 0$,因而 $\sigma(\alpha_1) = \alpha_1 \in \mathrm{Im}(\sigma)$,有 $V_1 \subseteq \mathrm{Im}(\sigma)$,故 $\mathrm{Im}(\sigma) = V_1$.

$\forall \alpha \in \mathrm{Ker}(\sigma)$，则 $\sigma(\alpha)=0$，而 $\alpha=\alpha_1+\alpha_2, \alpha_i \in V_i, i=1,2, \sigma(\alpha)=\alpha_1$，得 $\alpha=0$，即有 $\alpha=\alpha_2 \in V_2$，得 $\mathrm{Ker}(\sigma) \subseteq V_2$；反之 $\forall \alpha_2 \in V_2, \alpha_2=0+\alpha_2$，得 $\sigma(\alpha_2)=0$，即 $\alpha_2 \in \mathrm{Ker}(\sigma)$，得 $V_2 \subseteq \mathrm{Ker}(\sigma)$，故 $\mathrm{Ker}(\sigma)=V_2$.

由 σ 的定义，$\forall \xi \in V_1 \subseteq V, \xi_1=\xi_1+0$，有 $\sigma(\xi_1)=\xi_1$，即 $\sigma|_{V_1}(\xi_1)=\xi_1=\varepsilon\xi_1$，故 $\sigma|_{V_1}=\varepsilon$.

3) $\forall \alpha \in V, \sigma(\alpha) \in \mathrm{Im}(\sigma)=V_1$，因此 $\sigma(\sigma(\alpha))=\sigma(\alpha)$，即 $\sigma^2(\alpha)=\sigma(\alpha)$，故 $\sigma^2=\sigma$.

4) $\forall \xi \in V, \xi=\xi_1+\xi_2, \xi_1 \in \mathrm{Im}(\sigma), \xi_2 \in \mathrm{Ker}(\sigma)$，于是

$$(\varepsilon-\sigma)(\xi)=\xi-\sigma(\xi)=\xi_1+\xi_2-\sigma(\xi_1+\xi_2)=\xi_1+\xi_2-\xi_1=\xi_2,$$

因此 $\varepsilon-\sigma$ 是 V 的关于直和分解 $V=\mathrm{Im}(\sigma)\oplus\mathrm{Ker}(\sigma)$ 到 $\mathrm{Ker}(\sigma)$ 的射影.

5) 因为

$$\mathrm{Ker}(\sigma)=\{\alpha-\sigma(\alpha) \mid \alpha \in V\}=\mathrm{Im}(\varepsilon-\sigma),$$
$$\mathrm{Ker}(\varepsilon-\sigma)=\{\alpha \mid (\varepsilon-\sigma)\alpha=0, \alpha \in V\}=\mathrm{Im}(\sigma),$$

所以由 $V=\mathrm{Im}(\sigma)\oplus\mathrm{Ker}(\sigma)$ 得

$$V=\mathrm{Ker}(\varepsilon-\sigma)\oplus\mathrm{Ker}(\sigma)=\mathrm{Im}(\sigma)\oplus\mathrm{Im}(\varepsilon-\sigma)=\mathrm{Ker}(\varepsilon-\sigma)\oplus\mathrm{Im}(\varepsilon-\sigma).$$

6) 由 5) 知 $\mathrm{Im}(\varepsilon-\sigma)=\mathrm{Ker}(\sigma), \mathrm{Ker}(\varepsilon-\sigma)=\mathrm{Im}(\sigma)$，故 $\mathrm{Im}(\varepsilon-\sigma), \mathrm{Ker}(\varepsilon-\sigma)$ 为 σ-子空间.

7) 因为有 $(\varepsilon-1/2\sigma) \in L(V_n)$，使 $(\varepsilon+\sigma)(\varepsilon-1/2\sigma)=(\varepsilon-1/2\sigma)r(\varepsilon+\sigma)=\varepsilon$，故 $\varepsilon+\sigma$ 为可逆的线性变换.

8) 由 σ 的定义及 $V=\mathrm{Im}(\sigma)\oplus\mathrm{Ker}(\sigma)$ 知 $\mathrm{Im}(\sigma)$ 是 σ 的属于特征值 1 的特征子空间，$\mathrm{Ker}(\sigma)$ 是 σ 的属于特征值 0 的特征子空间，并且有 $\dim V=\dim\mathrm{Im}(\sigma)+\dim\mathrm{Ker}(\sigma)$，故 σ 可对角化.

3.6 向量空间的准素分解

前面我们已经看到，求相似矩阵的标准形式问题与向量空间关于一个线性变换的不变子空间分解有着密切的关系，因此，我们先从向量空间的分解开始.

3.6.1 最小多项式

设 V 是一个 n 维向量空间，V 的一切线性变换作成一个 n^2 维向量空间 $L(V)$. 因此，对于 V 的每一线性变换 σ，n^2+1 个线性变换 $\sigma^k, k=0,1,\cdots,n^2$，一定线性相关，于是存在一个非零多项式 $f(x)$，使得 $f(\sigma)=\theta$，这里 θ 表示 V 的零变换，即 σ 满足多项式 $f(x)$. 在 σ 所满足的一切非零多项式中，有一个次数最低的. 设 $p(x)$ 和 $q(x)$ 都是 σ 所满足的

次数最低的多项式，以 $p(x)$ 除 $q(x)$ 得 $q(x)=p(x)h(x)+r(x)$. 如果 $r(x)\neq 0$，那么 $r(x)$ 就是一个次数低于 $p(x)$ 的多项式，于是有 $r(\sigma)=q(\sigma)-p(\sigma)h(\sigma)=\theta$，这与 $p(x)$ 是 σ 所满足的次数最低的多项式这一事实矛盾. 因此 $r(x)=0$，从而 $p(x)\mid q(x)$.

同理，$q(x)\mid p(x)$. 这样，$p(x)$ 与 $q(x)$ 只能相差一个常数因子.

定义 3.6.1　我们把线性变换 σ 所满足的次数最低、首项系数是 1 的非零多项式叫做 σ 的**最小多项式**.

由以上的说明可知，n 维向量空间的每一线性变换都恰有一个最小多项式.

类似地可以定义一个 n 阶矩阵 A 的最小多项式. 一个最高次项系数是 1 的非零多项式 $p(x)$ 叫做矩阵 A 的最小多项式，如果 $p(A)=0$，而对于任意一个次数低于 $p(x)$ 的非零多项式 $q(x)$ 来说，$q(A)\neq 0$.

容易看出，相似矩阵有相同的最小多项式，并且向量空间 V 的一个线性变换 σ 的最小多项式就是 σ 关于 V 的任意基的矩阵的最小多项式.

例 3.6.1　矩阵 $A=\begin{pmatrix}2&0&0\\1&2&0\\0&0&2\end{pmatrix}$ 的最小多项式是 $(x-2)^2$. 因为 $A-2I=\begin{pmatrix}0&0&0\\1&0&0\\0&0&0\end{pmatrix}$，所以 $(A-2I)^2=0$，并且 A 不满足任何一次多项式.

例 3.6.2　设

$$A=\begin{pmatrix}A_1&O\\O&A_2\end{pmatrix},$$

是一个对角线分块矩阵，$p_1(x)$ 和 $p_2(x)$ 分别是矩阵的最小多项式. 如果 $p_1(x)$ 和 $p_2(x)$ 互素，那么 $p_1(x)p_2(x)$ 就是 A 的最小多项式.

事实上，设 $p(x)$ 是 A 的最小多项式. 由 $p(A)=0$ 得出 $p(A_1)=0, p(A_2)=0$，因此 $p(x)$ 可以同时被 A_1 的最小多项式 $p_1(x)$ 和 A_2 的最小多项式 $p_2(x)$ 整除，然而 $p_1(x)$ 与 $p_2(x)$ 互素，所以有 $p_1(x)p_2(x)$ 整除 $p(x)$；另一方面，我们有

$$p(A_1)p(A_2)=\begin{pmatrix}p_1(A_1)&O\\O&p_1(A_2)\end{pmatrix}\begin{pmatrix}p_2(A_1)&O\\O&p_2(A_2)\end{pmatrix}=O,$$

所以 $p_1(x)p_2(x)$ 可以被 A 的最小多项式 $p(x)$ 整除. 因此 $p(x)=p_1(x)p_2(x)$. 一般地，设

$$A=\begin{pmatrix}A_1&&&0\\&A_2&&\\&&\ddots&\\0&&&A_k\end{pmatrix}$$

是一个对角线分块矩阵，而 $p_1(x)p_2(x)$ 是 A_i 的最小多项式，$i=1,2,\cdots,k$，如果 $p_1(x),\cdots,p_k(x)$ 两两互素，那么最小多项式是 $p_1(x)\cdots p_k(x)$.

3.6.2 线性变换的准素分解

定理 3.6.1 设 σ 是 n 维向量空间 V 的一个线性变换，$p(x)$ 是 σ 的最小多项式. 令 $p(x)=(x-\lambda_1)^{\gamma_1}\cdots(x-\lambda_k)^{\gamma_k}$ 是 $p(x)$ 在复数域上的不可约因式分解，这里 $\lambda_1,\cdots,\lambda_k$ 是互不相同的复数，$r_1,\cdots,r_k$ 是正整数，又设

$$V_i=\mathrm{Ker}(\sigma-\lambda_i)^{r_i}=\{\xi\in V\mid(\sigma-\lambda_i)^{r_i}\xi=0\},\ i=1,2,\cdots,k,$$

那么：1) 每一子空间 V 都在 σ 之下不变；

2) $V=V_1\oplus\cdots\oplus V_k$；

3) 令 $\sigma_i=\sigma|_{V_i}$ 是 σ 在 V_i 上的限制，那么 σ_i 的最小多项式是 $(x-\lambda_i)^{r_i}$，$i=1,2,\cdots,k$.

证明 令 $f_i(x)=\dfrac{p(x)}{(x-\lambda_i)^{r_i}}=\prod\limits_{j\neq i}(x-\lambda_j)^{r_j}$，$i=1,2,\cdots,k$.

因为 $\lambda_1,\cdots,\lambda_k$ 是互素的多项式，于是存在多项式 $u_1(x),\cdots,u_k(x)$ 使得 $f_1(x)u_1(x)+\cdots+f_k(x)u_k(x)=1$，令 $g_i(x)=f_i(x)u_i(x)$，$i=1,2,\cdots,k$，那么 $g_1(x)+\cdots+g_k(x)=1$. 将线形变换 σ 代入这个等式得 $g_1(\sigma)+\cdots+g_k(\sigma)=\iota$，这里 ι 是 V 的单位变换，于是 V 的每一向量 ξ 可以写成 $\xi=g_1(\sigma)\xi+\cdots+g_k(\sigma)\xi$.

令 $W_1=g_i(\sigma)V=\mathrm{Im}(g_i(\sigma))$，$i=1,2,\cdots,k$，那么每一 $g_i(\sigma)$ 作为 σ 的多项式，都与 σ 可交换，因而每一 W_i 在 σ 之下不变.

等式 $\xi=g_1(\sigma)\xi+\cdots+g_k(\sigma)\xi$ 表明 $V=W_1+\cdots+W_k$.

我们证明 $W_i=V_i=\mathrm{Ker}(\sigma-\lambda_i)^{r_i}$，$i=1,2,\cdots,k$，并且上面的和是直和. 首先

$$(x-\lambda_i)^{r_i}g_i(x)=(x-\lambda_i)^{r_i}f_i(x)u_i(x)=p(x)u_i(x),$$

所以对于 V 中任意 ξ，$(x-\lambda_i)^{r_i}g_i(x)\in W_i$. 因此 $W_i\subseteq V_i$.

反过来，设 $\xi_i\in V_i$，如果 $j\neq i$，那么 $g_j(x)=f_j(x)u_j(x)$ 可以被 $(x-\lambda_i)^{r_i}$ 整除，从而 $g_j(\sigma)\xi=0$，于是由 $\xi=g_1(\sigma)\xi+\cdots+g_k(\sigma)\xi$ 得 $\xi_i=g_i(\sigma)\xi_i\in W_i$，$i=1,2,\cdots,k$，这样 $V_i=W_i$，$i=1,2,\cdots,k$，而且 $V=V_1+\cdots+V_k$，现在设 ξ 是 V 中任意向量，那么 $\xi=g_1(\sigma)\xi+\cdots+g_k(\sigma)\xi$. 如果 ξ 还可以表成 $\xi=\xi_1+\cdots+\xi_k$，$i=1,2,\cdots,k$，那么由上面的证明，我们有 $g_i(\sigma)\xi_i=\xi_i$，而 $g_i(\sigma)\xi_i=0$. 若 $j\neq i$，则 $g_i(\sigma)\xi=g_i(\sigma)\xi_i=\xi_i$，$i=1,2,\cdots,k$，所以 V 中每一向量被表成 $\xi=\xi_1+\cdots+\xi_i$ 的表示法是唯一的，从而是直和.

最后令 $\sigma_i=\sigma|_{V_i}$. 因为 $(\sigma-\lambda_i)^{r_i}V_i=\{0\}$，所以 $(x-\lambda_i)^{r_i}$ 能被 σ_i 的最小多项式整除，从而 σ_i 的最小多项式一定有 $(x-\lambda_i)^{s_i}$ 的形式，这里 $0<s_i\leqslant r_i$，$i=1,2,\cdots,k$. 而 $(x-\lambda_1)^{s_1},\cdots,(x-\lambda_k)^{s_k}$ 两两互素，所以 σ 的最小多项式是 $(x-\lambda_1)^{s_1}\cdots(x-\lambda_k)^{s_k}$. 由多项式不可约因式分解的唯一性得出 $s_i=r_i$，即 $(x-\lambda_i)^{r_i}$ 是 σ_i 的最小多项式.

定理 3.6.1 所给出的空间 V 的分解叫做 V 关于线性变换 σ 的**准素分解**.

推论 3.6.2　n 维向量空间 V 的一个线性变换可以对角化的充分且必要条件是它的最小多项式没有重根.

证明　设 σ 是 V 的一个线性变换，$p(x)$ 是 σ 的最小多项式. 在 $C[x]$ 里，$p(x)$ 可以分解成一次因式的幂的乘积：$p(x)=(x-\lambda_1)^{r_1}\cdots(x-\lambda_k)^{r_k}$，式中 $\lambda_1,\cdots,\lambda_k$ 是互不相同的复数. 由定理 3.6.1 知，$V=V_1\oplus\cdots\oplus V_k$，这里 $V_i=\mathrm{Ker}(\sigma-\lambda_i)^{r_i}$，$i=1,2,\cdots,k$.

如果 $p(x)$ 没有重根，那么 $r_1=\cdots=r_k=1$，从而 $V_i=\mathrm{Ker}(\sigma-\lambda_i)$ 是 σ 的属于本征值 λ_i 的本征子空间，因而 V 是 σ 的本征子空间的直和，所以 σ 可以对角化.

反过来，如果可以对角化，那么 V 可以分解成为 σ 的本征子空间的直和：

$$V=V_1\oplus\cdots\oplus V_k,i=1,2,\cdots,k,$$

其中 $V_i=\mathrm{Ker}(\sigma-\lambda_i)$，$i=1,2,\cdots,k$，$\lambda_1,\cdots,\lambda_k$ 是 σ 的一切互不相同的本征值. 因此对于任意 $\xi_i\in V_i$，我们有 $\sigma(\xi_i)=\lambda_i\xi_i$，即 $(\sigma-\lambda_i)\xi_i=0$，所以 σ 在 V_i 上的限制 $\sigma|_{V_i}$ 满足一次多项式 $x-\lambda_i$，它显然就是 $\sigma|_{V_i}$ 的最小多项式. 再由例 3.6.2，σ 的最小多项式是 $p(x)=(x-\lambda_1)\cdots(x-\lambda_k)$，它没有重根.

推论 3.6.3　设 σ 是 n 维向量空间 V 的一个线性变换，而 $V=V_1\oplus\cdots\oplus V_k$，是 V 关于 σ 的准素分解，这里 $V_i=\mathrm{Ker}(\sigma-\lambda_i)$，而 $\dim V_i=n_i$，$i=1,2,\cdots,k$，那么 σ 的特征多项式是

$$f(x)=(x-\lambda_1)^{n_1}\cdots(x-\lambda_k)^{n_k}.$$

由这个推论立即看出，线性变换 σ 的最小多项式的一切互不相同的根恰是 σ 的一切互不相同的本征值.

3.6.3　Cayley - Hamilton 定理

我们进一步证明，σ 的最小多项式整除 σ 的特征多项式，从而 σ 满足它的特征多项式，这就是所谓的凯莱—哈密顿(Cayley-Hamilton)定理. 为了证明这个定理，先证明一个简单的事实，这个事实在以后的讨论中还要用到.

引理 3.6.4　设 τ 是向量空间 V 的一个线性变换，$\xi\in V$. 如果存在一个正整数 s 使得 $\tau^s(\xi)=0$，而 $\tau^{s-1}(\xi)=0$，那么向量 $\xi,\tau(\xi),\cdots,\tau^{s-1}(\xi)$ 线性无关.

证明　设存在不全为零的数 $a_0,a_1,\cdots,a_{s-1}$，使 $a_0\xi+a_1\tau(\xi)+\cdots+a_{s-1}\tau^{s-1}(\xi)=0$. 令 a_i 是第一个不等于零的系数，$0\leqslant i\leqslant s-1$，那么上面的等式就是 $a_i\tau^i(\xi)+\cdots+a_{s-1}\tau^{s-1}(\xi)=0$，对等式两端施行线性变换 τ^{s-i-1} 得 $a_i\tau^{s-1}(\xi)=0$，因此必须 $\tau^{s-1}(\xi)=0$，这与题设矛盾.

定理 3.6.5(凯莱—哈密顿)　设 σ 是 n 维向量空间 V 的一个线性变换，$f(x)$ 是 σ 的特征多项式，那么 $f(\sigma)=0$.

证明 只需证明σ的最小多项式整除σ的特征多项式. 设σ的最小多项式是

$$p(x)=(x-\lambda_1)^{\gamma_1}\cdots(x-\lambda_k)^{\gamma_k},$$

$\lambda_1,\cdots,\lambda_k$是互不相同的复数,则$V$可以分解为在$\sigma$之下不变的子空间的直和,

$$V=V_1\oplus\cdots\oplus V_k,$$

并且σ在V上的限制σ_i的最小多项式是$(x-\lambda_i)^{\gamma_i}$. 而σ_i的特征多项式是$(x-\lambda_i)^{n_i}$,这里$n_i=\dim V_i$,而σ的特征多项式是$f(x)=(x-\lambda_1)^{n_1}\cdots(x-\lambda_k)^{n_k}$. 因为$(x-\lambda_i)^{\gamma_i}$是$\sigma_i$的最小多项式,所以总存在$V_i$的一个向量$\xi_i$使得$(\sigma_i-\lambda_i)^{r_i}\xi=0$,而$(\sigma_i-\lambda_i)^{r_i-1}\xi\neq0$. 于是由引理3.6.4知,向量$\xi_i,(\sigma_i-\lambda_i)\xi_i,\cdots,(\sigma_i-\lambda_i)^{r_i-1}\xi_i$线性无关,从而$r_i\leqslant n_i$. 这就是说,$(x-\lambda_i)^{\gamma_i}$整除$(x-\lambda_i)^{n_i}$,$i=1,2,\cdots,k$,因而$\sigma$的最小多项式$p(x)$整除$\sigma$的特征多项式$f(x)$.

与定理3.6.5平行,我们有

定理3.6.6 设A是一个n阶矩阵,而$f(x)$是A的特征多项式,那么$f(A)=0$.

3.6.4 线性变换的若当分解

让我们进一步观察定理3.6.1的那个准素分解,令V是复数域$\mathbf{C}$上一个n维向量空间,σ是V的一个线性变换,又设$p(x)=(x-\lambda_1)^{\gamma_1}\cdots(x-\lambda_k)^{\gamma_k}$是$\sigma$的最小多项式,则$\lambda_1,\cdots,\lambda_k$就是$\sigma$的一切互不相同的本征值. 而空间$V$可以分解成为子空间的直和:$V=V_1\oplus\cdots\oplus V_k$,这里$V_i=\mathrm{Ker}(\sigma-\lambda_i)^{\gamma_i}$,$i=1,2,\cdots,k$,它们都在$\sigma$之下不变.

令$\pi_i=g_i(\sigma)$. 因为π_i是σ的多项式,所以每一个子空间V_j在每一线性变换π_i之下不变,$1\leqslant i,j\leqslant k$,则这个线性变换$\pi_i$具有以下性质:

1) $\pi_1+\cdots+\pi_k=\iota$,这里ι表示空间V的单位变换;

2) π_i在V_i上的限制是V_i的单位变换;

3) 如果$i\neq j$,那么π_i在V_j上的限制是V_j的零变换.

这样,π_i把空间V的每一个向量ξ映成它在V_i中的分量ξ_i,我们把π_i叫做V在子空间V_i上的射影. 现在令$\delta=\lambda_1\pi_1+\cdots+\lambda_k\pi_k$,那么由2)和3),每一子空间$V_i$在$\delta$之下不变,并且$\delta$在子空间$V_i$上的限制是$V_i$的一个位似,位似系数是$\lambda_i$,因此,$\delta$是一个可以对角化的线性变换. 由于

$$\delta=\lambda_1\pi_1+\cdots+\lambda_k\pi_k=\lambda_1g_1(\sigma)+\cdots+\lambda_kg_k(\sigma),$$

所以δ是σ的一个多项式. 令$\nu=\sigma-\delta$,那么ν也是σ的一个多项式,所以每一子空间V_i也在ν之下不变. 对于$\xi_i\in V_i$,我们有

$$\nu^{r_i}(\xi_i)=(\sigma-\delta)^{\gamma_i}(\xi_i)=(\sigma-\lambda_i\pi_i)^{\gamma_i}=0,$$

令$r=\max(r_1,\cdots,r_k)$,那么对于任意$\xi\in V$,$\nu^r(\xi)=0$. 因此,$\nu^r=0$.

定义 3.6.2　设 ν 是向量空间 V 的一个线性变换. 如果存在一个正整数 r,使得 $\nu^r = 0$,那么就称 ν 是一个**幂零线性变换**,简称**幂零变换**.

这样,V 的每一线性变换 σ 都可以写成 $\sigma = \delta + \nu$,其中 δ 是一个可对角化的线性变换,ν 是一个幂零线性变换. 由于 δ 和 ν 都是 σ 的多项式,所以 $\delta\nu = \nu\delta$.

我们将要进一步证明,满足上述条件的 δ 和 ν 是由 σ 唯一确定的. 为此,先证明下面的

引理 3.6.7　令 δ_1 和 δ_2 是 $\mathbf{C}$ 上 n 维向量空间 V 的可对角化的线性变换,且 $\delta_1\delta_2 = \delta_2\delta_1$,那么存在 V 的一个基,使得 δ_1 和 δ_2 关于这同一个基的矩阵是对角形式.

证明　因为 δ_2 可以对角化,所以 V 可以分解为 δ_2 的本征子空间的直和:

$$V = V_{\lambda_1} \oplus \cdots \oplus V_{\lambda_s},$$

这里 $\lambda_1, \cdots, \lambda_s \in \mathbf{C}$ 是 δ_2 的互不相同的本征值.

$$V_{\lambda_i} = \{v \in V \mid \delta_2(v) = \lambda_i v\}, 1 \leqslant i \leqslant s.$$

每一 V_{λ_i} 自然在 δ_2 之下不变. 又因为 $\delta_1\delta_2 = \delta_2\delta_1$,所以 V_{λ_i} 也在 δ_1 之下不变. 这样,δ_1 在 V_{λ_i} 上的限制 $\delta_1|_{V_{\lambda_i}}$ 是 V_{λ_i} 的一个线性变换. 因为 δ_1 可以对角化,所以由推论3.6.2,δ_1 的最小多项式没有重根. 由例 3.6.2 可得,$\delta_1|_{V_{\lambda_i}}$ 在 V_{λ_i} 上的最小多项式整除 δ_1 的最小多项式,所以前者也没有重根. 再由推论 3.6.2 知,$\delta_1|_{V_{\lambda_i}}$ 是 V_{λ_i} 的一个可以对角化的线性变换,因此可以选取 V_{λ_i} 的一个基,使得 $\delta_1|_{V_{\lambda_i}}$ 关于这个基的矩阵是对角形式,而 $\delta_2|_{V_{\lambda_i}}$ 关于 V_{λ_i} 的任意基的矩阵都是单位矩阵的一个标量倍. 从每一个 $V_{\lambda_i}(1 \leqslant i \leqslant s)$ 中这样地选取一个基,拼起来成为 V 的基,那么 δ_1 和 δ_2 关于这个基的矩阵都是对角形式.

定理 3.6.8　设 σ 是复数域上 n 维向量空间 V 的一个线性变换,那么存在 V 的一个可对角化的线性变换 δ 和一个幂零线性变换 ν,使得,1) $\sigma = \delta + \nu$; 2) $\delta\nu = \nu\delta$.

可对角化的 δ 和幂零的 ν 由条件 1) 和 2) 唯一确定,并且它们都是 σ 的多项式.

与定理平行,我们有关于矩阵的若当分解,我们把满足条件 $N^r = 0$(对某一正整数 r) 的矩阵 N 叫做**幂零矩阵**。

定理 3.6.9　设 A 是复数域上一个 n 阶矩阵,那么存在一个可对角化的矩阵 D 和一个幂零矩阵 N,使得

1) $A = D + N$;

2) $DN = ND$.

可对角化的矩阵 D 和幂零矩阵 N 由条件 1) 和 2) 唯一确定,并且它们都是 A 的多项式.

类似地,我们把定理给出的关于矩阵 A 的分解 $A = D + N$ 叫做 A 的若当分解. D 叫做 A 的可对角化部分,N 叫做 A 的幂零部分.

第4章 欧氏空间与双线性函数

向量空间的概念是解析几何中二维、三维空间概念的推广，然而在一般向量空间里缺少几何度量的概念. 在这一章里，我们将分别在 $\mathbf{R}$ 和 $\mathbf{C}$ 上向量空间里定义“内积”，从而引入度量的概念，得到欧氏空间和酉空间，这样的向量空间在数学、物理学等许多领域都有重要的应用. 我们主要讨论欧氏空间和酉空间、正交变换、对称变换、双线性函数和二次型等内容。

4.1 欧氏空间

4.1.1 内积与欧氏空间定义

定义 4.1.1 设 V 是实数域 $\mathbf{R}$ 上的一个向量空间，在 V 上定义一个二元映射（实函数）$<,>: V\times V\to\mathbf{R}$，对 $\forall\xi,\eta,\zeta\in V, a\in\mathbf{R}$ 满足下列条件：

1) $<\xi,\eta>=<\eta,\xi>$，（对称性）

2) $<\xi+\eta,\zeta>=<\xi,\zeta>+<\eta,\zeta>$，（可加性）

3) $<a\xi,\eta>=a<\xi,\eta>$，（齐次性）

4) 当 $\xi\neq 0$ 时，$<\xi,\xi>\ >0$，（恒正性）

则称 $<\xi,\eta>$ 为 ξ 与 η 的**内积**，而 V 称为对这个内积来说的一个**欧几里德空间**.

定理 4.1.1 在一个欧氏空间 V 里，$\forall\xi,\eta\in V$，都有 $<\xi,\eta>^2\leqslant<\xi,\xi><\eta,\eta>$，当且仅当 ξ 与 η 线性相关时，上式等号才成立.

特别地，在欧氏空间 $\mathbf{R}^n$ 中，$\forall\alpha=(a_1,\cdots,a_n),\beta=(b_1,\cdots,b_n)$，有不等式

$$\left(\sum_{i=1}^n a_ib_i\right)^2\leqslant\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right)$$

成立. 通常称上式为柯西 — 施瓦兹不等式（简称柯 — 施不等式）.

例 4.1.1 设 $\alpha=(a_1,\cdots,a_n),\beta=(b_1,\cdots,b_n)\in\mathbf{R}^n$，则由柯 — 施不等式易得

$$\left|\sqrt{\sum_{i=1}^n a_i^2}-\sqrt{\sum_{i=1}^n b_i^2}\right|\leqslant\sqrt{\sum_{i=1}^n (a_i-b_i)^2}.$$

如果在 $\mathbf{R}^n$ 中具体给出向量 α,β 来,则由柯 — 施不等式可以建立许多不等式.

令 $\alpha=\left(\sqrt{2},\sqrt{\sqrt{2}},\cdots,\sqrt{\sqrt{\cdots\sqrt{2}}}\right)$,其中最后一项中含有 n 个根号,$\beta=(1,1,\cdots,1)$,

则 $<\alpha,\beta>=\sqrt{2}+\sqrt{\sqrt{2}}+\cdots+\sqrt{\sqrt{\cdots\sqrt{2}}}$,其中最后一项中含有 n 个根号.

$<\alpha,\alpha>=2+\sqrt{2}+\sqrt{\sqrt{2}}+\cdots+\sqrt{\sqrt{\cdots\sqrt{2}}}$,其中最后一项中含有 $n-1$ 个根号.

$<\beta,\beta>=1$.

于是由柯 — 施不等式即得如下不等式:

$$\left(\sqrt{2}+\sqrt{\sqrt{2}}+\cdots+\sqrt{\sqrt{\cdots\sqrt{2}}}\right)^2\leqslant n\left(2+\sqrt{2}+\cdots+\sqrt{\sqrt{\cdots\sqrt{2}}}\right).$$

例 4.1.2　设 α 为欧氏空间 V 的一个非零向量,$\alpha_1,\cdots,\alpha_m\in V$ 线性无关,若满足 1) $<\alpha_i,\alpha>>0,(i=1,\cdots,m)$;2) $<\alpha_i,\alpha_j>\leqslant 0,(i,j=1,\cdots,m,i\neq j)$.

证明　当 $x_1\alpha_1+\cdots+x_m\alpha_m=0$ 时,我们总可将这 m 个实数按正、负或零重新排序,使 $x_1,\cdots,x_r\geqslant 0;x_{r+1},\cdots,x_m\leqslant 0,(1\leqslant r\leqslant m)$.

令 $\xi=\sum\limits_{i=1}^{r}x_i\alpha_i=-\sum\limits_{j=r+1}^{m}x_j\alpha_j$,

则 $<\xi,\xi>=<\sum\limits_{i=1}^{r}x_i\alpha_i,-\sum\limits_{j=r+1}^{m}x_j\alpha_j>=-\sum\limits_{i=1}^{r}\sum\limits_{j=r+1}^{m}x_ix_j<\alpha_i,\alpha_j>\leqslant 0$. 由内积性质 $<\xi,\xi>\geqslant 0$,必有 $<\xi,\xi>=0$,即得 $\xi=0$. 进而 $\sum\limits_{i=1}^{r}x_i\alpha_i=0$, $\sum\limits_{j=r+1}^{m}x_j\alpha_j=0$,于是

$$<\sum_{i=1}^{r}x_i\alpha_i,\alpha>=\sum_{i=1}^{r}x_i<\alpha_i,\beta>=-\sum_{j=r+1}^{m}x_j\alpha_j=0,$$

$$<\sum_{j=r+1}^{m}x_j\alpha_j,\alpha>=\sum_{j=r+1}^{m}x_j<\alpha_j,\alpha>=0.$$

得 $x_i<\alpha_i,\alpha>=0,x_j<\alpha_j,\alpha>=0$,则 $x_i=x_j=0,i=1,\cdots,r,j=r+1,\cdots,m$. 因此,$\alpha_1,\cdots,\alpha_m$ 线性无关.

4.1.2　向量间的距离

定义 4.1.2　向量 ξ 的长度定义为 $|\xi|=\sqrt{<\xi,\xi>}$,因此有

1) $|\xi|\geqslant 0$, $\forall\xi\in V$.

2) $|a\xi|=a|\xi|$, $\forall\xi\in V,\forall a\in\mathbf{R}$.

3) $\zeta/|\zeta|$ 称为单位向量,$\forall\xi(\neq 0)\in V$.

定义 4.1.3　两个非零向量 ξ,η 的**夹角** θ 指的是 $\arccos\dfrac{<\xi,\eta>}{|\xi|\cdot|\eta|}$.

一般地,$\theta\in[0,\pi]$,特别当 $\theta=\pi/2$ 时,$<\xi,\eta>=0$,称 ξ 与 η **正交**.

定义 4.1.4　两个向量 ξ 与 η 的**距离**指的是 $|\xi-\eta|=d(\xi,\eta)$. 满足下面性质

1) 当 $\xi\neq\eta$ 时, $d(\xi,\eta)>0$;

2) $d(\xi,\eta)=d(\eta,\xi)$;

3) $d(\xi,\eta)\leqslant d(\xi,\zeta)+d(\eta,\zeta)$, 其中 $\forall\,\xi、\eta、\zeta\in V$.

定义 4.1.5　设 $V(R)$ 是 n 维欧氏空间, $\varepsilon_1,\cdots,\varepsilon_n$ 为 V 的一个基, 由基向量的内积构成的 n 阶实矩阵 A 称为基 $\{\varepsilon_1,\cdots,\varepsilon_n\}$ 的**度量矩阵**, 简称**度量阵**.

$$A=\begin{pmatrix}<\varepsilon_1,\varepsilon_1> & <\varepsilon_1,\varepsilon_2> & \cdots & <\varepsilon_1,\varepsilon_n>\\ <\varepsilon_2,\varepsilon_1> & <\varepsilon_2,\varepsilon_2> & \cdots & <\varepsilon_2,\varepsilon_n>\\ \vdots & \vdots & & \vdots\\ <\varepsilon_n,\varepsilon_1> & <\varepsilon_n,\varepsilon_2> & \cdots & <\varepsilon_n,\varepsilon_n>\end{pmatrix}$$

定理 4.1.2　(1) 度量阵是对称的; (2) 度量阵之间是合同的.

定理 4.1.3　设 $\{\alpha_1,\cdots,\alpha_n\}$ 为欧氏空间 V 的基, A 为基 $\{\alpha_1,\cdots,\alpha_n\}$ 的度量阵, 则

1) 度量阵 A 是可逆的;

2) 度量阵 A 是正定的;

3) 设 B 为基 $\{\beta_1,\cdots,\beta_n\}$ 的度量阵, 则 A 与 B 合同.

证明　1) 因为 $A=\begin{pmatrix}<\alpha_1,\alpha_1> & \cdots & <\alpha_1,\alpha_n>\\ \vdots & & \vdots\\ <\alpha_n,\alpha_1> & \cdots & <\alpha_n,\alpha_n>\end{pmatrix}$,

所以 A 可逆 $\Leftrightarrow |A|\neq 0 \Leftrightarrow AX=0$ 只有零解. 然而 n 元齐次方程组

$$AX=\begin{pmatrix}<\alpha_1,\alpha_1> & \cdots & <\alpha_1,\alpha_n>\\ \vdots & & \vdots\\ <\alpha_n,\alpha_1> & \cdots & <\alpha_n,\alpha_n>\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=0$$

只有零解 $\Leftrightarrow <\sum_{i=1}^{n}x_i\alpha_i,\alpha_j>=0\ (j=1,\cdots,n)$, 则只有 $x_1=x_2=\cdots=x_n=0$. 故由 $\{\alpha_1,\cdots,\alpha_n\}$ 是 V 的一个基知, $\alpha_1,\cdots,\alpha_n$ 线性无关, 即

若 $\sum_{i=1}^{n}x_i\alpha_i=0$, 则 $x_i=0, i=1,\cdots,n$, 亦即若 $<\sum_{i=1}^{n}x_i\alpha_i,\alpha_j>=0, j=1,\cdots,n$, 则 $x_i=0, i=1,\cdots,n$, 因此, n 元齐次方程组只有零解, 所以度量阵 A 是可逆的.

2) 由内积的性质知 $A'=A$, 即 A 为实对称阵. 对于 $\forall\,\alpha(\neq 0)\in V(R)$, $\alpha=\sum_{i=1}^{n}x_i\alpha_i$, $<\alpha,\alpha>>0$, 因而

$$<\alpha,\alpha>=<\sum_{i=1}^{n}x_i\alpha_i,\sum_{j=1}^{n}x_j\alpha_j>=\sum_{i=1}^{n}\sum_{j=1}^{n}x_ix_j<\alpha_i,\alpha_j>$$

$$=(x_1,\cdots,x_n)\begin{pmatrix}<\alpha_1,\alpha_1> & \cdots & <\alpha_1,\alpha_n>\\ \vdots & & \vdots\\ <\alpha_n,\alpha_1> & \cdots & <\alpha_n,\alpha_n>\end{pmatrix}\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=X'AX>0.$$

其中 $X'=(x_1,\cdots,x_n)'\neq 0$ 为 V 中任一向量 α 关于基 $\{\alpha_1,\cdots,\alpha_n\}$ 的坐标，并且是任一组不全为零的实数，故 A 为正定.

3) 设由基 $\{\alpha_1,\cdots,\alpha_n\}$ 到基 $\{\beta_1,\cdots,\beta_n\}$ 的过渡阵为 T，即 $(\beta_1,\cdots,\beta_n)=(\alpha_1,\cdots,\alpha_n)T$，其中 $T=(t_{ij})_n$，于是有

$$<\beta_i,\beta_j>=<\sum_{k=1}^{n}t_{ki}\alpha_k\sum_{h=1}^{n}t_{hj}\alpha_h>=\sum_{k=1}^{n}\sum_{h=1}^{n}t_{ki}t_{hj}<\alpha_k,\alpha_h>=\sum_{h=1}^{n}\Big(\sum_{k=1}^{n}t_{ki}<\alpha_k,\alpha_h>t_{hj}\Big).$$

上式右端恰为 $T'AT=T'\begin{pmatrix}<\alpha_1,\alpha_1> & \cdots & <\alpha_1,\alpha_n>\\ \vdots & & \vdots\\ <\alpha_n,\alpha_1> & \cdots & <\alpha_n,\alpha_n>\end{pmatrix}T$ 的第 i 行第 j 列元素，左端是 B 的第 j 行第 i 列元素，故 $B=T'AT$，因而 A 与 B 合同.

例 4.1.3　设 $\{\alpha_1,\cdots,\alpha_n\}$ 为 n 维欧氏空间 $V_n(R)$ 的一个基，求证：对任意 n 个实数 $b_1,\cdots,b_n$，恰有一个向量 $\alpha\in V_n(R)$，使 $<\alpha,\alpha_i>=b_i, i=1,\cdots,n$.

证明　由于基 $\{\alpha_1,\cdots,\alpha_n\}$ 的度量阵 A 可逆，以 A 为系数阵，以任意给定的 n 个实数 $b_1,\cdots,b_n$ 为常数项而构成的 n 元线性方程组 $A\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=\begin{pmatrix}b_1\\ \vdots\\ b_n\end{pmatrix}$，即 $\sum_{j=1}^{n}<\alpha_i,\alpha_j>x_j=b_i$，$i=1,\cdots,n$，必有唯一解，令为 $k_1,\cdots,k_n$，于是设 $\alpha=k_1\alpha_1+\cdots+k_n\alpha_n$，就是被 $k_1,\cdots,k_n$ 唯一确定的 V 中的向量，使

$$<\alpha,\alpha_i>=<\sum_{j=1}^{n}k_j\alpha_j,\alpha_i>=\sum_{j=1}^{n}k_j<\alpha_j,\alpha_j>=\sum_{j=1}^{n}<\alpha_i,\alpha_j>k_j=b_j.$$

4.1.3　标准正交基

定义 4.1.6　**标准正交基**指的是 $V(R)$ 的一个基 $\{\alpha_1,\cdots,\alpha_n\}$ 中每一基向量都是单位向量，且它们彼此正交，即 $<\alpha_i,\alpha_j>=0$，$<\alpha_i,\alpha_i>=1, i,j=1,2,\cdots,n$.

定理 4.1.4　欧氏空间 V 中一个正交组一定是线性无关组.

定理 4.1.5　对 n 维欧氏空间 V 中任意一组基 $\{\varepsilon_1,\cdots,\varepsilon_n\}$ 都可找到一组标准正交基 $\{\eta_1,\cdots,\eta_n\}$，使 $L(\varepsilon_1,\cdots,\varepsilon_n)=L(\eta_1,\cdots,\eta_n), i=1,2,\cdots,n$.

Schmidt 正交化方法：对欧氏空间 V 中任意一组线性无关的向量组 $\alpha_1,\cdots,\alpha_r$ 可以通过下面的公式求出一个正交组 $\beta_1,\cdots,\beta_r$，即

$$\beta_1 = \alpha_1,$$

$$\beta_k = \alpha_k - \frac{<\alpha_1,\beta_1>}{<\beta_1,\beta_1>}\beta_1 - \cdots - \frac{<\alpha_{k-1},\beta_{k-1}>}{<\beta_{k-1},\beta_{k-1}>}\beta_{k-1}, k = 2, \cdots, r.$$

对每一 β_i 进行单位化，即 $\gamma_i = \beta_i / |\beta_i|$, $i = 1, \cdots, r$，就得到一个标准正交组 $\{\gamma_1, \cdots, \gamma_r\}$.

定理 4.1.6 设 $\alpha_1, \cdots, \alpha_n$ 为 n 维欧氏空间 V 的一个标准正交基，下面命题等价：

1) $\forall \xi \in V, \xi = \sum_{i=1}^{n} a_i\alpha_i, \eta = \sum_{j=1}^{n} b_j\beta_j$，使 $<\xi,\eta> = \sum_{i=1}^{n} a_ib_i$.

2) $\forall \xi \in V, \xi = \sum_{i=1}^{n} a_i\alpha_i$，总有 $a_i = <\xi,\alpha_i>$, $i = 1,\cdots,n$.

3) 若 $(\varepsilon_1, \cdots, \varepsilon_n)$ 为 V 的一个标准正交基，且 $(\alpha_1, \cdots,\alpha_n) = (\varepsilon_1, \cdots, \varepsilon_n)A$，则 A 为正交阵.

4) 若 $\{\beta_1, \cdots, \beta_n\}$ 为标准正交基，且 $(\beta_1, \cdots, \beta_n) = (\alpha_1, \cdots, \alpha_n)A$，则 A 为正交阵.

5) 若 U 为正交阵，使 $(\beta_1, \cdots, \beta_n) = (\alpha_1, \cdots, \alpha_n)U$，则 $\{\beta_1, \cdots, \beta_n\}$ 为 V 的一个标准正交基.

6) 若 U 为正交阵，使 $(\alpha_1, \cdots, \alpha_n) = (\varepsilon_1, \cdots, \varepsilon_n)U$，则 $\{\varepsilon_1, \cdots, \varepsilon_n\}$ 为 V 的一个标准正交基.

把 $\mathbf{R}^n$ 的一个基 $\alpha_j = (a_{1j}, \cdots, a_{nj})'$, $j = 1, \cdots, n$，化为标准正交基的方法如下：

1) 令 $A = (\alpha_1, \cdots, \alpha_n)$，则 A 为实可逆阵，且 $A'A$ 为 n 阶正定阵；

2) 对 $A'A$ 施行合同交换，化 $A'A$ 为单位阵的同时，对单位阵 I_n 施行同样的列初等变换所得到的 n 阶可逆阵 T，使 $T'(A'A)T = I_n$.

3) 令 $(\alpha_1, \cdots, \alpha_n)T = (\beta_1, \cdots, \beta_n)$，其中 $\beta_j = (b_{1j}, \cdots, b_{nj})$, $j = 1, \cdots, n$，再令 $B = AT = (\beta_1, \cdots, \beta_n) = (b_{ij})_n$，则 $B'B = T'(A'A)T = I$. 于是

$$<\beta_i, \beta_j> = \sum_{k=1}^{n} b_{ki}b_{kj} = \begin{cases} 1, i = j; \\ 0, i \neq j. \end{cases}$$

这样由 $(\alpha_1, \cdots, \alpha_n)$ 得到的 $(\beta_1, \cdots, \beta_n)$ 就是 $\mathbf{R}^n$ 的一个标准正交基.

4.1.4 欧氏空间的同构

定义 4.1.7 设 $W \leqslant V$，则称子空间 $W^{\perp} = \{\xi \in V \mid <\xi,W> = 0\}$ 为 W 的**正交补**.

定理 4.1.7 设 W 为欧氏空间 V 的一个有限维子空间，则 $V = W \oplus W^{\perp}$，即 V 可分解为它的有限维子空间与其正交补的直和。

定理 4.1.8 设 V_1 与 V_2 都是 n 维欧氏空间 V 的子空间，且 $\dim V_1 < \dim V_2$，则 V_2 中必有一非零向量 β 与 V_1 正交.

证明 欲证 V_1 中的非零向量与 V_2 正交，只须证 V_2 中必有非零向量属于 V_1 的正交补 $V_1^{\perp}$ 即可. 对此令 $\dim V_1 = r_1$，则由 $V = V_1 \oplus V_1^{\perp}$ 知 $\dim V_1^{\perp} = n - r_1$，这样 V_2 与 $V_1^{\perp}$，或 $V_2 \cap V_1^{\perp} = \{0\}$，或 $V_2 \cap V_1^{\perp} \neq \{0\}$，两者必居其一. 若 $V_2 \cap V_1^{\perp} = \{0\}$，则由 $\dim V_2 = r_2 r_1$，推得 $\dim(V_2 + V_1^{\perp}) = \dim V_2 + \dim V_1^{\perp} = n + (r_2 - r_1) > n$，这不可能，故必有 $V_2 \cap V_1^{\perp} \neq \{0\}$，但 $V_2 \cap V_1^{\perp} \neq \varnothing$，因此有 $\beta(\neq 0) \in V_2 \cap V_1^{\perp}$，即 $\beta \in V_2$，且 $\beta \in V_1^{\perp}$，所以 V_2 中非零向量 β 正交于 V_1.

定义 4.1.8 设 V 与 W 为两个欧氏空间，如果

1) $f: V \to W$ 为一个同构映射；

2) $\forall \xi, \eta \in V$，都有 $< f(\xi), f(\eta) > = < \xi, \eta >$，

则称 V 与 W 是**同构**的，记为 $V \cong W$.

定理 4.1.9 两个有限维欧氏空间同构的充分必要条件是它们的维数相等，进而任意 n 维欧氏空间都与 $\mathbf{R}^n$ 同构.

例 4.1.4 设 $\{\alpha_1, \cdots, \alpha_m\}$，$\{\beta_1, \cdots, \beta_m\}$ 为 n 维欧氏空间 V 的两组向量. 求证：若 $< \alpha_i, \alpha_j > = < \beta_i, \beta_j >, i, j = 1, \cdots, m$，则

1) $L(\alpha_1, \cdots, \alpha_m)$ 与 $L(\beta_1, \cdots, \beta_m)$ 同构；

2) 构造一个 V 的正交变换 σ，使 $\sigma(\alpha_i) = \beta_i (i = 1, \cdots, m)$.

证明 1) 欲证 $L(\alpha_1, \cdots, \alpha_m)$ 与 $L(\beta_1, \cdots, \beta_m)$ 同构，只须证它们的维数相等，等价于这两向量组的秩相等，为此，令 $\{\alpha_1, \cdots, \alpha_r\}$ 为 $\{\alpha_1, \cdots, \alpha_m\}$ 的一个极大无关组，r 为其秩，$1 \leqslant r \leqslant m$，则

$$| G(\alpha_1, \cdots, \alpha_r) | = \begin{vmatrix} < \alpha_1, \alpha_1 > & \cdots & < \alpha_1, \alpha_r > \\ \vdots & & \vdots \\ < \alpha_r, \alpha_1 > & \cdots & < \alpha_r, \alpha_r > \end{vmatrix} \neq 0,$$

由于 $< \alpha_i, \alpha_j > = < \beta_i, \beta_j >, i, j = 1, \cdots, m$，推得 $| G(\beta_1, \cdots, \beta_r) | \neq 0$. 进而 $\{\beta_1, \cdots, \beta_r\}$ 线性无关，因此 $\dim L(\beta_1, \cdots, \beta_m) \geqslant r$，即 $\dim L(\beta_1, \cdots, \beta_m) \geqslant \dim L(\alpha_1, \cdots, \alpha_m)$.

同理可证 $\dim L(\alpha_1, \cdots, \alpha_m\} \geqslant \dim L(\beta_1, \cdots, \beta_m)$.

所以 $L(\alpha_1, \cdots, \alpha_m\}$ 与 $L\{\beta_1, \cdots, \beta_m\}$ 的维数相等，它们同构.

2) 设 $V_1 = L(\alpha_1, \cdots, \alpha_m)$，$V_2 = L(\beta_1, \cdots, \beta_m)$，则 $V = V_1 \oplus V_1^{\perp} = V_2 \oplus V_2^{\perp}$.

由 V_1 与 V_2 同构，即 V_1 与 V_2 间有一同构映射，令为 f，使 $f(\alpha_i) = \beta_i, i = 1, \cdots, m$. 因此由 $\dim V_1 = \dim V_2$ 有 $\dim V_1^{\perp} = \dim V_2^{\perp}$，即 $V_1^{\perp}$ 与 $V_2^{\perp}$ 间还有一同构映射，令为 g，于是令 $\sigma: \xi \mapsto f(\xi) + g(\xi)$，其中 $\forall \xi \in V = V_1 \oplus V_2, \xi = \xi_1 + \xi_1'$.

易证 σ 为 $V_n(R)$ 的一个线性变换，并且保持内积，故 σ 为一个正交变换. 同时，$\forall \alpha_i \in V_1 (i = 1, \cdots, m)$，$\alpha_i = \alpha_i + 0$，有 $\sigma(\alpha_i) = f(\alpha_i) + g(0) = \beta_i, i = 1, \cdots, m$，所以 σ 即为所求.

反之,由正交变换 σ,使 $\sigma(\alpha_i)=\beta_i,i=1,\cdots,m$,易推得 $<\alpha_i,\alpha_j>=<\beta_i,\beta_j>,i,j=1,\cdots,m$,成立.因此 $<\alpha_i,\alpha_j>=<\beta_i,\beta_j>,i,j=1,\cdots,m$,不仅是构造 V 的一个正交变换 σ,使 $\sigma(\alpha_i)=\beta_i,i=1,\cdots,m$,的充分条件,而且还是必要条件.

4.2 正交变换和对称变换

4.2.1 正交变换

定义 4.2.1 **正交变换**指的是保持内积不变的线性变换,即 $\sigma\in L(V)$,对 $\forall\xi,\eta\in V$,都有 $<\sigma(\xi),\sigma(\eta)>=<\xi,\eta>$.

定理 4.2.1 设 σ 为 n 维欧氏空间 V 的一个线性变换,则下面诸命题彼此等价:

(1) σ 是正交变换;

(2) $\forall\xi\in V$,$|\sigma(\xi)|=|\xi|$,即 σ 保持向量长度不变;

(3) 若$\{\varepsilon_1,\cdots,\varepsilon_n\}$ 为一个标准正交基,则 $\sigma(\varepsilon_1),\cdots,\sigma(\varepsilon_n)$ 仍是标准正交基;

(4) σ 关于任一个标准正交基的矩阵是正交阵.

定理 4.2.2 设 V 为 n 维欧氏空间,σ 为 V 的一个正交变换,令 $V_1=\{\alpha\in V\mid\sigma(\alpha)=\alpha\}$,$V_2=\{\alpha-\sigma(\alpha)\mid\alpha\in V\}$ 为 V 的子空间,则 $V=V_1\oplus V_2$.

证明 $\forall\alpha\in V_1\cap V_2$,于是 $\alpha=\sigma(\alpha)$,$\alpha=(\varepsilon-\sigma)\beta$,则

$$<\alpha,\alpha>=<\alpha,(\beta-\beta)>=<\alpha,\beta>-<\alpha,\sigma(\beta)>=<\alpha,\beta>-<\sigma(\alpha),\sigma(\beta)>=0,$$

故 $\alpha=0$,即 $V_1\cap V_2=\{0\}$.

$$V_1=\{\alpha\in V\mid\sigma(\alpha)=\alpha\}=\{\alpha\in V\mid(\varepsilon-\sigma)\alpha=0\}=\mathrm{Ker}(\varepsilon-\sigma);$$

$$V_2=\{(\varepsilon-\sigma)\alpha\mid\alpha\in V\}=\mathrm{Im}(\varepsilon-\sigma),$$

其中 ε 为恒等变换,因为 $\dim\mathrm{Im}(\varepsilon-\sigma)+\dim\mathrm{Ker}(\varepsilon-\sigma)=n$,所以由 $V_1\cap V_2=\{0\}$ 及 $V_1+V_2\subseteq V$ 得 $V_1+V_2=V$,进而 $V=V_1\oplus V_2$.

例 4.2.1 设 σ 为欧氏空间 V 上的正交变换,且 $\sigma^m=\varepsilon$($m>1$,ε 为恒等变换),令 $V_\sigma=\{\xi\in V\mid\sigma(\xi)=\xi\}$,称为 σ 的固定点子空间,$V_\sigma^\perp$ 为 V_σ 的正交补空间.求证:

1) $V_\sigma^\perp$ 是 σ -子空间;

2) $\forall\xi\in V$,记$\bar{\xi}=\dfrac{1}{m}\sum\limits_{i=0}^{m-1}\sigma^i(\xi)$,则$\bar{\xi}\in V_\sigma$;

3) $\forall\xi\in V=V_\sigma\oplus V_\sigma^\perp$,$\xi$ 的分解式 $\xi=\xi_1+\xi_2$($\xi_1\in V_\sigma$, $\xi_2\in V_\sigma^\perp$)中 $\xi_1=\bar{\xi}$.

证明 1) 略;2) 略;

3) $\forall\xi\in V=V_\sigma\oplus V_{1/\sigma}$,则 $\xi=\xi_1+\xi_2$,其中 $\xi_1\in V_\sigma$,$\xi_2\in V_{1/\sigma}$,于是

$$\sigma(\xi)=\sigma(\xi_1)+\sigma(\xi_2)=\xi_1+\sigma(\xi_2),$$

$$\sigma^2(\xi) = \sigma(\xi_1) + \sigma^2(\xi_2) = \xi_1 + \sigma^2(\xi_2),$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\sigma^{m-1}(\xi) = \xi_1 + \sigma^{m-1}(\xi_2),$$

$$\xi = \xi_1 + \sigma^m(\xi_2) = \xi_1 + \xi_2,$$

进而 $\xi + \sigma(\xi) + \cdots + \sigma^{m-1}(\xi) = m\xi_1 + \xi_2 + \sigma(\xi_2) + \cdots + \sigma^{m-1}(\xi_2)$，由 2) 即得 $m\bar{\xi} = m\xi_1 + m\xi_2^*$，其中令 $m\xi_2^* = \xi_2 + \sigma(\xi_2) + \cdots + \sigma^{m-1}(\xi_2)$，则$\bar{\xi} = \xi_1 + \xi_2^*$. 由于$\bar{\xi}, \xi_1 \in V_\sigma$，得 $\xi_2^* \in V_\sigma$. 又由 1) 知 $V_{1/\sigma}$ 在σ之下不变，而 $\xi_2 \in V_{1/\sigma}$，因此 $\sigma^i(\xi_2) \in V_{1/\sigma}, i = 0, 1, \cdots, m-1$，所以 $\xi_2^* \in V_{1/\sigma}$，故 $\xi_2^* \in V_\sigma \cap V_{1/\sigma} = \{0\}$，得$\xi_2^* = 0$，则$\bar{\xi} = \xi_1$.

定理 4.2.3　设 V 为 n 维欧氏空间，$\sigma, \tau \in L(V)$，$\forall \alpha \in V$，有 $< \sigma(\alpha), \sigma(\alpha) > = < \tau(\alpha), \tau(\alpha) >$，求证：

1) $\mathrm{Im}(\sigma)$ 与 $\mathrm{Im}(\tau)$ 同构；

2) V 中存在一个正交变换 δ，则 $\delta\sigma = \tau$.

4.2.2　对称变换

定义 4.2.2　设 $\sigma \in L(V)$，$\forall \xi, \eta \in V$，有 $< \sigma(\xi), \eta > = < \xi, \sigma(\eta) >$，则称 σ 为**对称变换**.

定理 4.2.4　设 σ 为 n 维欧氏空间 V 的一个线性变换，则下面诸命题彼此等价：

(1) σ 为对称变换；

(2) σ 关于任一个标准正交基的矩阵是对称阵；

(3) σ 关于某一标准正交基的矩阵是对角形阵.

定理 4.2.5　n 阶实对称阵 A 的特征值全为实数，并且属于 A 的不同特征值的特征向量彼此正交.

定理 4.2.6　对于任一个 n 阶实对称阵 A，都存在一个 n 阶正交阵 U，使 $U'AU = U^{-1}AU$ 为对角形，且其主对角线上的元素为 A 的全部特征值.

例 4.2.2　设 σ 为 n 维实向量空间 V 上的线性变换. 证明：能在 V 上引入内积，使得 σ 为对称变换的 $\Leftrightarrow \sigma$ 有 n 个线性无关的特征向量.

证明　*必要性*　若在 V 上引入内积，则 V 是一个欧氏空间，并且σ为对称变换，于是 σ 关于 V 的一个标准正交基 $e_1, \cdots, e_n$ 的矩阵 A 为实对称阵，因此有正交阵 U，使 $U'AU = \mathrm{diag}(\lambda_1, \cdots, \lambda_n)$，其中 $\lambda_1, \cdots, \lambda_n$ 为 A 的全部特征值. 令：

$$(r_1, \cdots, r_n) = (e_1, \cdots, e_n)U,$$

则 $r_1, \cdots, r_n$ 为 V 的一个标准正交基. 这样 σ 关于基 $\{r_1, \cdots, r_n\}$ 的矩阵为 $\mathrm{diag}(\lambda_1, \cdots, \lambda_n)$，即 $\sigma(r_1, \cdots, r_n) = (e_1, \cdots, e_n)\mathrm{diag}(\lambda_1, \cdots, \lambda_n)$，从而有 $\sigma(r_i) = \lambda r_i, i = 1, 2, \cdots, n$. 故

$(r_1,\cdots,r_n)$ 为 σ 的线性无关的特征向量.

充分性　设 $\alpha_1,\cdots,\alpha_n$ 为 σ 的 n 个线性无关的特征向量,因此它们构成 V 的一个基.任意 $\xi,\eta\in V$,有 $\xi=\sum a_i\alpha_i,\eta=\sum b_i\alpha_i$,

对此我们定义一个 $V\times V\to R$ 的二元映射:$<\xi,\eta>=\sum a_ib_i$,这样定义的二元映射满足内积四条公理,于是在 V 上引入了内积,由 $\sigma(\alpha_i)=\lambda_i\alpha_i,i=1,2,\cdots,n$,有

$$<\xi,\sigma(\eta)>=<\sum a_i\alpha_i,\sigma(\sum b_i\alpha_i)>=<\sum a_i\alpha_i,\sum b_i\lambda_i\alpha_i>=\sum a_ib_i\lambda_i;$$

$$<\sigma(\xi),\eta>=<\sigma(\sum a_i\alpha_i),\sum b_i\alpha_i>=<\sum a_i\lambda_i\alpha_i,\sum b_i\alpha_i>=\sum a_ib_i\lambda_i,$$

故 σ 为对称变换.

例 4.2.3　设 σ 为 n 维欧氏空间 V 的一个反对称变换,即 $\forall\alpha,\beta\in V$,都有 $<\sigma(\alpha),\beta>=-<\alpha,\sigma(\beta)>$,求证:

1) 线性变换 $\sigma\pm\varepsilon$ 均为满秩的,其中 ε 为恒等变换;

2) 线性变换 $\tau=(\sigma-\varepsilon)(\sigma+\varepsilon)^{-1}$ 为正交变换.

证明　设 $e_1,\cdots,e_n$ 为 V 的一个标准正交基,而 A 与 I 分别为 σ 与 ε 关于这个基的矩阵,则 $\sigma\pm\varepsilon$ 在这个基下的矩阵为 $A\pm I$. 因此欲证上述结论,只须证它们关于这个基下的矩阵的可逆性与正交性.

1) 因 σ 为反对称变换,故 A 为反对称阵. 易知 A 的特征根只能是零或纯虚数,因而 $|1I-A|\neq0$,即 $|1A-I|\neq0$,于是 $\sigma-\varepsilon$ 为满秩的. 又因 $(A+I)'=A'+I=I-A$,进而 $A+I$ 可逆,所以 $(\sigma+\varepsilon)$ 也是满秩的.

2) 由于 $\tau=(\sigma-\varepsilon)(\sigma+\varepsilon)^{-1}$ 在基 $e_1,\cdots,e_n$ 下的矩阵为 $T=(A-I)(A+I)^{-1}$,因此只须证 T 为正交矩阵.

$$\begin{aligned}T'T&=((A-I)(A+I)^{-1})'(A-I)(A+I)^{-1}\\&=((A+I)^{-1})'(A-I)'(A-I)(A+I)^{-1}\\&=(I-A)^{-1}(A+I)(I-A)(A+I)^{-1}\\&=(I-A)^{-1}(I-A)(A+I)(A+I)^{-1}=I,\end{aligned}$$

故 τ 为正交阵,于是 $\tau=(\sigma-\varepsilon)(\sigma+\varepsilon)^{-1}$ 为正交变换.

4.3 酉 空 间

4.3.1 酉空间的定义

定义 4.3.1　设 V 是复数域 $\mathbf{C}$ 上的向量空间,在 V 上定义一个二元映射 $(,)$: $V\times$

$V \to \mathbf{C}$ 满足下列条件：

(1) $(\xi,\eta)=\overline{(\eta,\xi)}$，其中$\overline{(\eta,\xi)}$是$(\eta,\xi)$的共轭复数；

(2) $(k\xi,\eta)=k(\xi,\eta)$；

(3) $(\xi+\eta,\zeta)=(\xi,\zeta)+(\eta,\zeta)$；

(4) (ξ,ξ)为非负实数，且$(\xi,\xi)=0 \Leftrightarrow \xi=0$.

其中 $\forall \xi,\eta,\zeta \in V, k\in \mathbf{C}$，则称$(\xi,\eta)$为向量$\xi$与$\eta$的内积，而$V$称为对这个内积而言的一个**酉空间**，记为$V(\mathbf{C})$.

内积的表示，设$\{\alpha_1,\cdots,\alpha_n\}$为$n$维酉空间$V$的一个基，$\forall \xi,\eta\in V$，

$$\xi=\sum_{i=1}^{n}a_i\alpha_i,\eta=\sum_{j=1}^{n}b_j\alpha_j,\text{则}(\xi,\eta)=\left(\sum_{i=1}^{n}a_i\alpha_i,\sum_{j=1}^{n}b_j\alpha_j\right)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_i\overline{b_j}(\alpha_i,\alpha_j).$$

定义 4.3.2 $\sqrt{(\alpha,\alpha)}$称为向量α的**长度**，记为$|\alpha|$.

定理 4.3.1 对于$\forall \xi,\eta\in V(\mathbf{C})$，$|(\xi,\eta)|\leqslant|\xi||\eta|$，等号成立的充要条件是$\xi,\eta$线性相关.酉空间中内积$(\xi,\eta)$一般是复数，向量间不易定义夹角，因此酉空间没有夹角概念.

定义 4.3.3 当$(\xi,\eta)=0$时，称ξ与η**正交或垂直**，记为$\xi\perp\eta$.

在n维酉空间中，有正交基和标准正交基，因此任意一组线性无关的向量可通过Schmidt正交化方法化为正交组，并扩充为V的一个标准正交基.

定义 4.3.4 设V是酉空间，W为V的有限维子空间，

$$W^{\perp}=\{\xi,\eta\in V(\mathbf{C})\mid(\xi,\eta)=0,\forall\eta\in V\}\text{ 称为 }W\text{ 的\textbf{正交补}}.$$

定理 4.3.2 若$W^{\perp}$为W的正交补，则$V=W\oplus W^{\perp}$.

定义 4.3.5 设$A\in M_n(\mathbf{C})$，若$\overline{A'}A=A\overline{A'}=I$，即$A^{-1}=\overline{A'}$，则称$A$为**酉矩阵**，且$\|A\|=1$，即$A$的行列式的绝对值等于1.

定理 4.3.3 n维酉空间的两组标准正交基的过渡阵是酉矩阵.

定理 4.3.4 1) 两个酉矩阵的乘积仍是酉矩阵，酉矩阵的逆和转置都是酉矩阵；

2) 酉矩阵的行列式的模等于1；

3) 酉矩阵的特征根的模等于1.

例 4.3.1 设σ是n维酉空间V的一个线性变换，则下面四个命题彼此等价：

1) σ是酉变换；

2) σ是V的同构映射；

3) 若$\{\alpha_1,\cdots,\alpha_n\}$为$V$的一个标准正交基，则$\sigma(\alpha_1),\cdots,\sigma(\alpha_n)$也是$V$的标准正交基；

4) σ关于任一标准正交基的矩阵是酉矩阵.

证明 1)⇒2) 由σ是酉变换及同构的定义，知σ是V的同构映射；2)⇒1) 是显然的.

1)⇒3) 当$i=j$时$(\alpha_i,\alpha_j)=1$；当$i\neq j$时$(\alpha_i,\alpha_j)=0$，$i,j=1,\cdots,n$，而σ为酉

变换，故

$$(\sigma(\alpha_i),\sigma(\alpha_j)) = \begin{cases} 1, i=j, \\ 0, i\neq j, \end{cases} \quad i,j=1,\cdots,n,$$

即$\{\sigma(\alpha_1),\cdots,\sigma(\alpha_n)\}$为$V$的一个标准正交基.

反之，3)⇒1)亦然.事实上，由$\{\alpha_1,\cdots,\alpha_n\}$与$\{\sigma(\alpha_1),\cdots,\sigma(\alpha_n)\}$为$V$的两个标准正交基，则由$\xi=\sum x_i\alpha_i,\eta=\sum y_j\alpha_j,\sigma(\xi)=\sum x_i\sigma(\alpha_i),\sigma(\eta)=\sum y_j\sigma(\alpha_j)$可得，$<\xi,\eta>=\sum x_iy_i=<\sigma(\xi),\sigma(\eta)>$，故$\sigma$是酉变换，即1)⇔3).

3)⇒4) 设σ关于一标准正交基$\{\beta_1,\cdots,\beta_n\}$的矩阵为$A$，则$\{\sigma(\beta_1),\cdots,\sigma(\beta_n)\}$也是$V$的一个标准正交基，于是$A$可看作由标准正交基$\{\beta_1,\cdots,\beta_n\}$到标准正交基$\{\sigma(\beta_1),\cdots,\sigma(\beta_n)\}$的过渡阵，因此$A$是酉矩阵.反之，由4)知$A$是酉矩阵，即$A\overline{A'}=\overline{A'}A=I$，于是由$(\sigma(\alpha_1),\cdots,\sigma(\alpha_n))=(\alpha_1,\cdots,\alpha_n)A$，经计算得

$$(\sigma(\alpha_i),\sigma(\alpha_j)) = \begin{cases} 1, i=j, \\ 0, i\neq j, \end{cases} \quad i,j=1,\cdots,n,$$

故$\{\sigma(\alpha_1),\cdots,\sigma(\alpha_n)\}$为$V$的一个标准正交基，所以3)⇔4).

例 4.3.2 设A为n阶正交阵，必存在一个n阶酉矩阵U，使$U'AU$为对角形阵.

证明 设V为一个n维酉空间，$\{\alpha_1,\cdots,\alpha_n\}$为一个标准正交基，令$\sigma$是$A$关于基$\{\alpha_1,\cdots,\alpha_n\}$下确定的线性变换，由于$A$为$n$阶正交阵，必有$A\overline{A'}=\overline{A'}A=I$，即$A$为$n$阶酉矩阵，因此$\sigma$为一个酉变换.由于酉空间的任一酉变换，都存在由它的特征向量构成的标准正交基，使σ在此基下的矩阵为对角形，令$(\beta_1,\cdots,\beta_n)=(\alpha_1,\cdots,\alpha_n)U$，$U$为$n$阶酉矩阵，且

$$\sigma(\beta_1,\cdots,\beta_n)=(\beta_1,\cdots,\beta_n)\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix},\text{使 } U^{-1}AU=\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}.$$

4.3.2 酉变换和对称变换

定义 4.3.6 酉空间V的线性变换σ，若$\forall\xi,\eta\in V$，满足$(\sigma(\xi),\sigma(\eta))=(\xi,\eta)$，则称$\sigma$为$V$的一个**酉变换**.

定理 4.3.5 σ是n维酉空间V的一个线性变换，则

1) σ是酉变换；

2) 若$\{\varepsilon_1,\cdots,\varepsilon_n\}$为一个标准正交基，则$\sigma(\varepsilon_1),\cdots,\sigma(\varepsilon_n)$仍是标准正交基；

3) σ关于任一个标准正交基的矩阵是酉矩阵.

定理 4.3.6　酉变换 σ 存在一个由其特征向量构成的标准正交基,使 σ 关于这个基的矩阵为对角形.

定义 4.3.7　酉空间 V 的线性变换 σ,若 $\xi,\eta\in V$,满足 $(\sigma(\xi),\eta)=(\xi,\sigma(\eta))$,则称 σ 为 V 的一个**对称变换**.

定义 4.3.8　令 $H\in M_n(\mathbf{C})$,若 $\overline{H'}=H$,则称 H 为 **Hermite 矩阵**.

定理 4.3.7　设 σ 为 n 维酉空间 V 的一个对称变换,则

1) σ 的本征值全为实数;

2) σ 的属于不同本征值的本征向量彼此正交;

3) 存在 V 的一个标准正交基,使得 σ 关于这个基的矩阵是 n 阶 Hermite 矩阵.

定理 4.3.8　n 维酉空间 V 的一个对称变换,则下面诸命题彼此等价:

1) σ 为对称变换;

2) σ 关于任一个标准正交基的矩阵是 Hermite 矩阵.

定理 4.3.9　H 是一个 n 阶 Hermite 矩阵, 存在 n 阶酉矩阵 U,使得

$$\overline{U'}HU=U^{-1}HU$$

是实对角形,即任意 n 阶 Hermite 矩阵都"酉相似" 于一个实对角形阵.

例 4.3.3　设 A 为 n 阶复阵,若满足 $A\overline{A'}=\overline{A'}A$,则称 A 为复正规阵,试证: n 阶复阵 A 为复正规阵的充要条件是酉相似于对角阵,即存在 n 阶酉矩阵 U,使 $\overline{U'}AU$ 为对角形阵.

证明　*充分性*　由 $\overline{U'}AU$ 为对角形知 $\overline{U'}AU$ 亦是复正规阵,于是有 $\overline{U'}A\,\overline{A'}U=\overline{U'}\,\overline{A'}AU$,故 $A\overline{A'}=\overline{A'}A$,即 A 为复正规阵.

必要性　令 λ_1 为 n 阶复正规阵 A 的一个特征根,η_1 为属于特征根 λ_1 的单位特征向量,于是有一个以 η_1 为第一列的酉矩阵 $U_1=(\eta_1,\cdots,\eta_n)$, 使

$$\overline{U'_1}AU_1=\begin{pmatrix}\lambda_1 & a_2 & \cdots & a_n\\ 0 & & & \\ \vdots & & A_1 & \\ 0 & & & \end{pmatrix}$$

由于 A 是复正规的,知 $\overline{U'_1}AU_1$ 也是副正规的,于是由

$$\begin{pmatrix}\overline{\lambda_1} & 0 & \cdots & 0\\ \overline{a_2} & & & \\ \vdots & & \overline{A'_1} & \\ \overline{a_n} & & & \end{pmatrix}\begin{pmatrix}\lambda_1 & \cdots & a_n\\ \vdots & A_1 & \\ 0 & & \end{pmatrix}=\begin{pmatrix}\lambda_1 & \cdots & a_n\\ \vdots & A_1 & \\ 0 & & \end{pmatrix}\begin{pmatrix}\overline{\lambda_1} & 0 & \cdots & 0\\ \overline{a_2} & & & \\ \vdots & & \overline{A'_1} & \\ \overline{a_n} & & & \end{pmatrix}$$

推出 $a_2 = \cdots = a_n = 0$，且 A_1 是 $n-1$ 阶复正规阵.

这样对 n 归纳证明，显然当 $n=1$ 时成立；

假定当 $n-1$ 时成立，于是令

$$U_0 = \begin{pmatrix} 1 & \\ & U_2 \end{pmatrix}, U = U_1 U_0,$$

则

$$\overline{U'}AU = \overline{U'_0}\,\overline{U'_1}A\,U_1U_0$$

$$= \begin{pmatrix} 1 & \\ & \overline{U'_2} \end{pmatrix}\begin{pmatrix} \lambda_1 & \\ & A_1 \end{pmatrix}\begin{pmatrix} 1 & \\ & U_2 \end{pmatrix} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}.$$

其中 U 是酉矩阵 U_1 与酉矩阵 U_0 之积，仍是酉矩阵.

4.4 双线性函数

4.4.1 线性函数的定义及性质

定义 4.4.1 设 $V(F)$ 为数域 F 上的向量空间，f 是一个映射，若 f 满足

1) $f(\alpha+\beta) = f(\alpha) + f(\beta)$ （可加性）

2) $f(k\alpha) = kf(\alpha)$ （齐次性）

其中 $\forall \alpha, \beta \in V(F), \forall k \in F$，则称 f 为 V 上一个**线性函数**.

线性函数有如下简单性质：

1) $f(0) = 0$；$f(-\alpha) = -f(\alpha)$；

2) $\beta = \sum_{i=1}^{r} k_i \alpha_i, f(\beta) = \sum_{i=1}^{r} k_i f(\alpha_i)$.

定理 4.4.1 设 $\alpha_1, \cdots, \alpha_n$ 为 $V_n(F)$ 的一个基，$a_1, \cdots, a_n$ 为 F 中任意 n 个数，则存在唯一的 V 上线性函数 f，使 $f(\alpha_i) = a_i, i = 1, \cdots, n$.

例 4.4.1 证明：对于 n 维向量空间 V，$\forall \alpha \neq \beta \in V$，必存在 V 上的线性函数 f，使 $f(\alpha) \neq f(\beta)$，即对 $\forall 0 \neq \xi \in V$，必有 $f \in V^*$，使 $f(\xi) \neq 0$.

证明 令 $e_1, \cdots, e_n$ 为 V 的一个基，而 $f_1, \cdots, f_n$ 为 V 的关于 $e_1, \cdots, e_n$ 的对偶基，则 $\forall \xi \in V$，有 $\xi = k_1 e_1 + \cdots + k_n e_n$，使 $f_i(\xi) = f_i(k_1 e_1 + \cdots + k_n e_n) = k_i, i = 1, \cdots, n$.

因此，由 $\xi \neq 0$ 可知，至少有一个 $k_r \neq 0, 1 \leqslant r \leqslant n$，所以 $f_r(\xi) = k_r \neq 0$，于是设 $f = f_r$ 即为所求. 亦即对 $\forall \alpha, \beta \in V$，且 $\alpha \neq \beta$，即 $\alpha - \beta \neq 0$，故 $f(\alpha - \beta) \neq 0$，进而 $f(\alpha) \neq f(\beta)$.

例 4.4.2　设 $V_n(F)$ 为 F 上 n 维向量空间，$g_1,\cdots,g_m \in V^*$，$1 \leqslant m < n$. 证明：必有向量 $\alpha \in V(\alpha \neq 0)$，使 $g_i(\alpha) = 0,\ i = 1,\cdots,m$.

证明　设 $\alpha_1,\cdots,\alpha_n$ 为 V 的基，令 $g_i(\alpha_j) = a_{ij}$，$i = 1,\cdots,m,\ j = 1,\cdots,n$. 于是对于 V 中某一非零向量 $\alpha = k_1\alpha_1 + \cdots + k_n\alpha_n$，倘若 $g_i(\alpha) = k_1 g_i(\alpha_1) + \cdots + k_n g_i(\alpha_n) = a_{i1}k_1 + \cdots + a_{in}k_n = 0$，则非零向量 α 的坐标：$k_1,\cdots,k_n$ 是齐次线性方程组的一个非零解，

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = 0, \\ \qquad\cdots\cdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = 0. \end{cases}$$

因此，由 $m < n$ 知，齐次线性方程组必有非零解，进而存在非零向量 α，使 $g_i(\alpha) = 0$.

4.4.2　对偶空间

定义 4.4.2　设 $L(V,F)$ 为数域 F 上 n 维向量空间 $V_n(F)$ 的全体线性函数的集合，在 $L(V,F)$ 中定义

加法：$(f+g)(\alpha) = f(\alpha) + g(\alpha)$，$\forall \alpha \in V$，$\forall f,g \in L(V,F)$；

数量乘法：$(kf)(\alpha) = k(f(\alpha))$，$\forall k \in F$，$\forall \alpha \in V$，$\forall f \in L(V,F)$.

$L(V,F)$ 对上面的运算构成的向量空间称为 V 的**对偶空间**，记为 V^*，$V^* = L(V,F)$.

定义 4.4.3　设 $\varepsilon_1,\cdots,\varepsilon_n$ 为 $V_n(F)$ 的一个基，由 $f_i = \begin{cases} 1, i = j, \\ 0, i \neq j, \end{cases}$ $i,j = 1,\cdots,n$，构造的 V 上 n 个线性函数，便成为 V^* 的一个基，并称为 $\{\varepsilon_1,\cdots,\varepsilon_n\}$ 的**对偶基**.

定理 4.4.2　$\dim V(F) = \dim L(V,F) = \dim V^*$.

定理 4.4.3　设 $\{\alpha_1,\cdots,\alpha_n\}$ 与 $\{\beta_1,\cdots,\beta_n\}$ 为 $V_n(F)$ 的两个基，它们的对偶基分别为 $\{f_1,\cdots,f_n\}$ 与 $\{g_1,\cdots,g_n\}$，若由 $\{\alpha_1,\cdots,\alpha_n\}$ 到 $\{\beta_1,\cdots,\beta_n\}$ 的过渡阵为 A，则由 $f_1,\cdots,f_n$ 到 $g_1,\cdots,g_n$ 的过渡阵为 $(A')^{-1}$，即若 $(\beta_1,\cdots,\beta_n) = (\alpha_1,\cdots,\alpha_n)A$，则 $(g_1,\cdots,g_n) = (f_1,\cdots,f_n)(A')^{-1}$.

定理 4.4.4　设 V 为一向量空间，V^{**} 是 V 的对偶空间 V^* 的对偶空间，V 到 V^{**} 的映射 $x \to x^{**}$ 是一个同构映射，从而 V 与 V^{**} 同构.

例 4.4.3　设 $f_1,\cdots,f_r$ 为数域 F 上 n 维向量空间 $V_n(F)$ 的 r 个线性函数. 令 $W = \{\alpha \in V \mid f_i(\alpha) = 0, 1 \leqslant i \leqslant r\}$，求证：

1) W 为 V 的一个子空间，并称 W 为线性函数 $f_1,\cdots,f_n$ 的零化子空间；

2) V 的任一子空间 W 均可视为某些线性函数的零化子空间.

证明　1) 证略.

2) 若 $W = \{0\}$，则 W 显然是 V 上任意线性函数的零化子空间；若 $W = V$，则取零函

数. 现令 W 为 V 的非平凡子空间, $\alpha_1, \cdots, \alpha_r$ 为 W 的一个基, 并将其扩为 V 的一个基 $\alpha_1, \cdots, \alpha_r, \alpha_{r+1}, \cdots, \alpha_n$, 由此基得到它们的一对偶基 $f_1, \cdots, f_r, f_{r+1}, \cdots, f_n$, 于是 $f_i(\alpha_j) = 0, i = r+1, \cdots, n, j = 1, \cdots, r$.

$$\forall \beta \in W, \beta = \sum_{j=1}^{r} a_j \alpha_j, \text{则 } f_i(\beta) = \sum_{j=1}^{r} a_j f_i(\alpha_j) = 0, i = r+1, \cdots, n.$$

从而 $\beta \in \{\alpha \in V \mid f_i(\alpha) = 0, r+1 \leqslant i \leqslant n\}$, 即得 $W \subseteq \{\alpha \in V \mid f_i(\alpha) = 0, r+1 \leqslant i \leqslant n\}$;

反之, $\forall \beta \in \{\alpha \in V \mid f_i(\alpha) = 0, r+1 \leqslant i \leqslant n\}$, 由 $\beta \in V$, 有 $\beta = \sum_{i=1}^{n} f_i(\beta)\alpha_i = \sum_{j=1}^{r} f_j(\beta)\alpha_j$, 即 $W \supseteq \{\alpha \in V \mid f_i(\alpha) = 0, r+1 \leqslant i \leqslant n\}$, 故 $W = \{\alpha \in V \mid f_i(\alpha) = 0, r+1 \leqslant i \leqslant n\}$ 是 V 的一个线性函数 $f_{r+1}, \cdots, f_n$ 的零化子空间.

4.4.3 双线性函数

定义 4.4.4 设 $f(\alpha, \beta)$ 是 $V(F)$ 上的一个二元函数, 即任意 $\alpha, \beta \in V(F)$, 根据 f 都唯一确定 F 中数 $f(\alpha, \beta)$, 若满足

(1) 设 $f(\alpha, k_1\beta_1 + k_2\beta_2) = k_1 f(\alpha, \beta_1) + k_2 f(\alpha, \beta_2)$;

(2) $f(k_1\alpha_1 + k_2\alpha_2, \beta) = k_1 f(\alpha_1, \beta) + k_2 f(\alpha_2, \beta)$,

其中 $\forall k_1, k_2 \in F, \forall \alpha, \alpha_1, \alpha_2, \beta, \beta_1, \beta_2 \in V$, 则称 $f(\alpha, \beta)$ 为 V 上的一个**双线性函数**.

例 4.4.4 设 $f(\alpha, \beta)$ 为 n 维向量空间 $V_n(F)$ 的一个双线性函数, W 为 V 的 m 维子空间, 则

1) $W^{\perp} = \{\beta \in V \mid f_i(\alpha, \beta) = 0, \forall \alpha \in W\}$ 为 W 的子空间. 特别地, $V^{\perp}$ 称为 V 的右迷向子空间.

2) 若 $\beta \in V, \beta \notin W^{\perp}$, 则 $U = \{\alpha \in V \mid f_i(\alpha) = 0\}$ 是 W 的 $n-1$ 维子空间.

证明 1) 证略.

2) 由于 $\beta \notin W^{\perp}$, 故有 $\alpha \in W$, 使得 $f(\alpha, \beta) \neq 0$, 令 $f(\alpha, \beta) = k, \alpha_0 = k^{-1}\alpha$, 则 $f(\alpha_0, \beta) = 1$, $\forall \xi \in W$, 再令 $f(\xi, \beta) = h$, 则 $\xi = (\xi, h\alpha_0) + h\alpha_0$, 从而由 $f(\xi - h\alpha_0, \beta) = f(\xi, \beta) - hf(\alpha_0, \beta) = 0$ 知 $\xi - h\alpha_0 \in U$, 由此可见, W 中任一向量 ξ 均可写成 U 中一个元素 $\xi - h\alpha_0$ 与 $h\alpha_0$ 的和, 即 $W = U + L(\alpha_0)$, 其中 $L(\alpha_0)$ 是由 α_0 生成的一维子空间, 倘若 $L(\alpha_0)$ 中任一元 $a\alpha_0 \in U$, 则由 $0 = f(a\alpha_0, \beta) = af(\alpha_0, \beta) = a$ 知, $U \cap L(\alpha_0) = (0)$, 即得 $W = U \oplus L(\alpha_0)$, 因此 $\dim U = m - 1$.

定义 4.4.5 设 $f(\alpha, \beta)$ 是 $V_n(F)$ 上一个双线性函数, $\{\varepsilon_1, \cdots, \varepsilon_n\}$ 是 V 的一个基, 则矩阵

$$A=\begin{pmatrix} f(\varepsilon_1,\varepsilon_1) & f(\varepsilon_1,\varepsilon_2) & \cdots & f(\varepsilon_1,\varepsilon_n) \\ f(\varepsilon_2,\varepsilon_1) & f(\varepsilon_2,\varepsilon_2) & \cdots & f(\varepsilon_2,\varepsilon_n) \\ \vdots & \vdots & & \vdots \\ f(\varepsilon_n,\varepsilon_1) & f(\varepsilon_n,\varepsilon_2) & \cdots & f(\varepsilon_n,\varepsilon_n) \end{pmatrix}$$

称为 $f(\alpha,\beta)$ 在$\{\varepsilon_1,\cdots,\varepsilon_n\}$下的**度量矩阵**.

在不同基下,$V_n(F)$ 上同一个双线性函数 $f(\alpha,\beta)$ 的度量矩阵彼此是合同的,即若 A 为 $f(\alpha,\beta)$ 在基$\{\alpha_i \mid i=1,\cdots,n\}$下的度量矩阵,$B$ 为 $f(\alpha,\beta)$ 在基$\{\beta_i \mid i=1,\cdots,n\}$下的度量矩阵,而 T 是由基$\{\alpha_1,\cdots,\alpha_n\}$到基$\{\beta_1,\cdots,\beta_n\}$的过渡阵,则 $B=T^{-1}AT$.

例 4.4.5 设 $f(\alpha,\beta)$ 为 n 维向量空间 V 上一个双线性函数,若它的度量阵 A 的秩是 r,则称 $f(\alpha,\beta)$ 是秩 r 的双线性函数,证明非零双线性函数 $f(\alpha,\beta)$ 可分解为两个线性函数之积的充要条件是它的秩为 1.

证明 设 $e_1,\cdots,e_n$ 为 V 的一个基,任意 α、$\beta\in V$,有

$$\alpha=\sum_{i=1}^{n}x_ie_i,\beta=\sum_{i=1}^{n}y_ie_i,$$

令 A 为 $f(\alpha,\beta)$ 的度量矩阵,由 $f(\alpha,\beta)\neq 0$,有 $A\neq 0$. 若 $f(\alpha,\beta)=X'AY=f(\alpha,\beta)$,其中 $X'=(x_1,\cdots,x_n)$,$Y'=(y_1,\cdots,y_n)$, $f_1(\alpha)=\sum_{i=1}^{n}a_ix_i$,$f_2(\beta)=\sum_{i=1}^{n}b_iy_i$,则

$f(\alpha,\beta)=X'\left(\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}(b_1\cdots b_n)\right)Y$,即 $A=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}(b_1\cdots b_n)$,则 $R(A)=1$,故 $f(\alpha,\beta)$ 的秩为 1;反之,$f(\alpha,\beta)$ 的度量矩阵 A 的秩为 1,则必有 $A=\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}(b_1\cdots b_n)$,因此 $f(\alpha,\beta)=$

$X'AY=X'\begin{pmatrix}a_1\\ \vdots\\ a_n\end{pmatrix}(b_1\cdots b_n)Y=\left(\sum_{i=1}^{n}a_ix_i\right)\left(\sum_{i=1}^{n}b_iy_i\right)$,于是令 $f_1(\alpha)=\sum_{i=1}^{n}a_ix_i$,$f_2(\beta)=\sum_{i=1}^{n}b_iy_i$,即得 $f(\alpha,\beta)=f_1(\alpha)f_2(\beta)$,故非零双线性函数 $f(\alpha,\beta)$ 可表为两个线性函数 $f_1(\alpha)$ 与 $f_2(\beta)$ 的乘积.

定义 4.4.6 设 $f(\alpha,\beta)$ 为 $V_n(F)$ 的一个双线性函数,若对 $\forall\beta\in V$,都有 $f(\alpha,\beta)=0$,可推出 $\alpha=0$,则 f 称为**非退化的**.

定理 4.4.5 双线性函数 $f(\alpha,\beta)$ 非退化的充要条件是 $f(\alpha,\beta)$ 在基$\{\alpha_1,\cdots,\alpha_n\}$下的度量矩阵是可逆的.

4.4.3 对称双线性函数

定义 4.4.7 设 $f(\alpha,\beta)$ 是 $V(F)$ 上一个双线性函数，若对 $\forall \alpha,\beta \in V(F)$，都有 $f(\alpha,\beta) = f(\beta,\alpha)$，则称 $f(\alpha,\beta)$ 为**对称双线性函数**.

又 $\forall \alpha,\beta \in V(F)$，有 $f(\alpha,\beta) = -f(\beta,\alpha)$，则称 $f(\alpha,\beta)$ 为**反对称双线性函数**.

定理 4.4.6 双线性函数是对称的 $\Leftrightarrow$ 它在任一基下的度量矩阵是对称矩阵.

定理 4.4.7 设 $f(\alpha,\beta)$ 是 $V_n(F)$ 上一个对称双线性函数，则存在 V 的一个基 $(\varepsilon_1, \cdots, \varepsilon_n)$ 使 $f(\alpha,\beta)$ 在这基下的度量矩阵为对角形阵，因此数域 F 上每一个 n 阶对称阵 A 都与 n 阶对角阵合同，即存在一个 n 阶可逆阵 P，使 $P'AP = \operatorname{diag}(c_1, \cdots, c_n)$.

定义 4.4.8 设 $f(\alpha,\beta)$ 为 $V(F)$ 上对称双线性函数，当 $\alpha=\beta$ 时，V 上的函数 $f(\alpha,\alpha)$ 称为与 $f(\alpha,\beta)$ 关联的**二次齐次函数**.

定理 4.4.8 设 V 是复数域上的向量空间，维数 $n \geqslant 2$，$f(\alpha, \beta)$ 是 V 上一个对称双线性函数，

1) 证明，V 中有非零向量 ξ 使 $f(\xi, \xi) = 0$；

2) 若 $f(\alpha, \beta)$ 非退化，必有线性无关向量 ξ, η 满足 $f(\zeta, \eta) = 1, f(\xi, \xi) = f(\eta, \eta) = 0$.

证明 1) 由 $\dim V = n \geqslant 2$，α,β 为 V 的两个线性无关向量，令 $f(\alpha,\alpha) = a, f(\beta,\beta) = b, f(\alpha,\beta) = c$，若 a,b 有一为零，则结论成立；若 a,b 皆非零，则令 $\xi = \alpha + \lambda\beta$，且 $\xi \neq 0$，于是由 $f(\xi,\xi) = f(\alpha,\alpha) + 2\lambda f(\alpha,\beta) + \lambda^2 f(\beta,\beta)$ 知，当 λ 为 $bx^2 + 2cx + a$ 的根时，就使 $f(\xi,\xi) = 0$，故只要取 $\lambda = -\varepsilon \pm \sqrt{c^2 - ab}/b$ 即可.

2) 由(1) 有 $\xi \neq 0$，使 $f(\xi,\xi) = 0$，又 $f(\alpha,\beta)$ 非退化，则存在 $u \in V$，使 $f(\xi,u) = b \neq 0$. 令 $v = u/b$，则 $f(\xi,v) = 1$，若 $f(v,v) = 0$，则令 $\eta = v$，即为所求.

若 $f(v,v) = a \neq 0$，则令 $\eta = v - a/2\xi$，于是

$$f(\eta,\eta) = f(v,v) - 2(a/2)f(\xi,v) + a^2/4f(\xi,\xi) = 0;$$

同时

$$f(\xi,\eta) = f(\xi,v - a/2\xi) = f(\xi,v) - a/2f(\xi,\xi) = 1.$$

例 4.4.6 设 f 是 n 维空间 V 的对称双线性函数，$\sigma \in L(V)$，对 $\forall \xi, \eta \in V$，$f(\sigma(\xi)), \sigma((\eta)) = f(\zeta, \eta)$. 设 f 关于 V 的某一基的矩阵是 A，σ 关于同一基的矩阵是 S，问 S 应满足什么条件?

解 令 $\alpha_1,\cdots,\alpha_n$ 为 V 一基，$A = (a_{ij}), a_{ij} = f(\alpha_i,\alpha_j), i,j = 1,\cdots,n$.

$$(\sigma(\alpha_1),\cdots,\sigma(\alpha_n)) = (\alpha_1,\cdots,\alpha_n)S, S = (b_{ij}),$$

于是 $\sigma(\alpha_j) = \sum\limits_{k=1}^{n} b_{kj}\alpha_k$，因此 $a_{ij} = f(\alpha_i,\alpha_j) = f(\sigma(\alpha_i),\sigma(\alpha_j)) = f\left(\sum\limits_{k=1}^{n} b_{ki}\alpha_k, \sum\limits_{l=1}^{n} b_{lj}\alpha_l\right)$

$$= \sum_{k=1}^{n}\sum_{l=1}^{n} b_{ki}b_{lj}f(\alpha_k,\alpha_l) = (b_{1i},\cdots,b_{ni})A\begin{pmatrix} b_{1j} \\ \vdots \\ b_{rj} \end{pmatrix},$$

故有 $A = S'AS$.

反之,若 $A = S'AS$, $\forall \xi, \eta \in V$,有 $f(\sigma(\xi), \sigma(\eta)) = f(\xi,\eta)$,那么对基 $\alpha_1,\cdots,\alpha_n$,有

$$f(\alpha_i, \alpha_j) = a_{ij} = (b_{1i}, \cdots, b_{ni})A\begin{pmatrix} b_{1j} \\ \vdots \\ b_{rj} \end{pmatrix} = \sum_{k=1}^{n}\sum_{l=1}^{n} b_{ki}b_{lj}f(\alpha_k,\alpha_l) = f(\sum_{k=1}^{n} b_{ki}\alpha_k, \sum_{l=1}^{n} b_{lj}\alpha_l)$$

$$= f(\sigma(\alpha_i), \sigma(\alpha_j)),$$

故 S 满足 $A = S'AS$.

4.5　二次型与正定矩阵的应用

4.5.1　二次型及二次型的矩阵

定义 4.5.1　设 $f(\alpha,\beta)$ 为 $V_n(F)$ 上一个对称双线性函数,$\{\alpha_2,\cdots,\alpha_n\}$ 为 $V_n(F)$ 上的一个基,A 为 $f(\alpha,\beta)$ 在基$\{\alpha_2,\cdots,\alpha_n\}$ 下的度量矩阵. $\forall \alpha \in V(F)$, $\alpha = \sum_{i=1}^{n} x_i\alpha_i$,有

$$f(\alpha,\alpha) = \sum_{i=1}^{n}\sum_{j=1}^{n} f(\alpha_i,\alpha_j)x_ix_j = (x_1,\cdots,x_n)A\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix},$$

其中 $A' = A$,上式右端是 F 上 n 个文字 $x_1,\cdots,x_n$ 的一个二次齐次多项式,称它为 F 上 n 个文字的**二次型**,简称 n 元二次型. 一般形式为

$$q(x_1,\cdots,x_n) = a_{11}x_1^2 + \cdots + a_{nn}x_n^2 + 2a_{12}x_1x_2 + \cdots + 2a_{(n-1)n}x_{n-1}x_n$$
$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_ix_j = X'AX,$$

其中 $A = (a_{ij})_{n\times n} = A'$, $i, j = 1, \cdots, n$, $X = (x_1, \cdots, x_n)'$,称 A 为 **n 元二次型 q 的矩阵**.

二次型的秩指的是它的矩阵的秩.

例 4.5.1　设实二次型 $f(x_1,\cdots,x_n) = \sum_{j=1}^{n}\sum_{i=1}^{n} a_{ij}x_ix_j = X'AX$, $(A' = A)$ 是满秩的,则称二次型 $q(x_1,\cdots,x_n) = \sum_{j=1}^{n}\sum_{k=1}^{n} \frac{A_{ij}}{|A|}x_ix_j$ 为 $f(x_1,\cdots,x_n)$ 的逆式,求证:满秩的二次型与它的逆式有相同的符号差,因而有相同的惯性指数.

证明 因为 $f(x_1,\cdots,x_n)$ 为满秩的,因此它的矩阵是可逆的,即 A^{-1} 存在,对此令非退化的变量替换为 $X=A^{-1}Y$,则

$$f(x_1,\cdots,x_n)=(A^{-1}Y)'A(A^{-1}Y)=Y'AY$$

$$=Y'\frac{A}{|A|}Y=\frac{1}{|A|}(y_1\cdots y_n)A^*\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$$

$$=\sum_{j=1}^{n}\sum_{i=1}^{n}\frac{A_{ij}}{|A|}y_iy_j=q(y_1,\cdots,y_n),$$

在非退化变换之下,二次型的正、负惯性指数不变,所以符号差不变.因此,二次型与其逆式有相同的符号差与惯性指数.

定义 4.5.2 设 A, B 是数域 F 上两个 n 阶方阵,如果存在 F 上可逆矩阵 P,使得 $P'AP=B$,那么称 B 与 A **合同**.矩阵合同是一个等价关系.

定理 4.5.1 数域 F 上两个二次型等价的充分必要条件是它们的矩阵合同.

等价二次型有相同的秩.因此,数域 F 上 n 元二次型

$$q(x_1,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=X'AX,$$

总可通过可逆变量替换 $\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=P\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$ 把它化为二次型 $c_1y_1^2+c_2y_2^2+\cdots+c_ny_n^2$.

4.5.2 复二次型与实二次型

定理 4.5.2 实数域 **R** 上 n 元二次型的规范形是由正惯性指数 p 唯一确定.因此通常称 p 为二次型的惯性指标,此命题称为**惯性定律**.

定理 4.5.3 复数域 **C** 上两个 n 元二次型等价当且仅当它们的秩相等;

实数域 **R** 上两个 n 元二次型等价当且仅当它们的秩相等且符号差也相等;

定义 4.5.3 称实二次型 $x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_r^2$ 为**规范形**(**或典范形**).

称其中正平方项的个数 p 为**正惯性指数**,负平方项的个数 $r-p$ 为负惯性指数.而它们的差 $p-(r-p)$ 称为二次型的**符号差**.

定理 4.5.4 实数域上每一 n 元二次型都等价于规范形 $x_1^2+\cdots+x_p^2-x_{p+1}^2-\cdots-x_r^2$,其中 p 为正惯性指数,r 为秩;平行地有 **R** 上 n 阶对称阵 A 合同于

$$\begin{pmatrix}I_p & & \\ & -I_{r-p} & \\ & & 0\end{pmatrix}$$

其中 I_p 与 I_{r-p} 分别为 p 阶与 $r-p$ 阶单位阵，r 为 A 的秩.

例 4.5.2 设 $q(x_1,\cdots,x_n)=X'AX$ 是秩为 n 的实二次型，则存在 $\mathbf{R}^n$ 的 $1/2(n-|s|)$ 维子空间 V_1（s 为符号差），使对任一向量 $(x_1,\cdots,x_n)\in V_1$，有 $q(x_1,\cdots,x_n)=0$.

证明 设 $q(x_1,\cdots,x_n)$ 的符号差 $s=2p-n$，于是存在可逆变量替换 $X=PY$，使

$$q(x_1,\cdots,x_n)=y_1^2+\cdots+y_p^2-y_{p+1}^2-\cdots-y_n^2,$$

由 $1/2(n-|s|)=\begin{cases}n-p,s\geqslant 0,\\ p,s<0,\end{cases}$ 以下分两种情形讨论之：

1）当 $s\geqslant 0$ 时，$1/2(n-s)=n-p$，令 $e'_j=(0,\cdots,0,\overset{(j)}{1},0,\cdots,0,\overset{(p+j)}{1},0,\cdots,0)$，$j=1,\cdots,n-p$，这 $\mathbf{R}^n$ 中 $n-p$ 个向量显然是线性无关的，并且 $q(e_j)=0,j=1,\cdots,n-p$，下面由 $e_1,\cdots,e_{n-p}$ 构造出 $X_1,\cdots,X_{n-p}$，这 $n-p$ 个线性无关向量所生成的 $n-p$ 维子空间 V_1 即为所求.

事实上，令 $X_j=Pe_j$，$j=1,\cdots,n-p$，显然 $X_j\in\mathbf{R}^n,j=1,\cdots,n-p$，并且 $X_1,\cdots,X_{n-p}$ 线性无关的，假如

$$\sum_{j=1}^{n-p}k_jX_j=0,\sum_{j=1}^{n-p}k_jPe_j=0,$$

则由 P 的可逆性及 $e_1,\cdots,e_n$ 的线性无关性得 $k_1=\cdots=k_{n-p}=0$，于是令 $V_1=L(X_1,\cdots,X_{n-p})$ 就是 $n-p=1/2(n-|s|)$ 维的 $\mathbf{R}^n$ 的子空间. 同时，对 V_1 中任一向量 $X=(x_1,\cdots,x_n)$，有 $X=\sum\limits_{i=1}^{n-p}a_iX_i$，由可逆矩阵决定的替换 $X=PY$，得

$$Y=P^{-1}X=P^{-1}\sum_{i=1}^{n-p}a_iX_i=\sum_{i=1}^{n-p}a_ie_i=(a_1,\cdots,a_{n-p},0,\cdots,0,\cdots,a_1,\cdots,a_{n-p}),$$

进而有

$$q(x_1,\cdots,x_n)=a_1^2+\cdots+a_{n-p}^2\cdots-a_1^2-\cdots-a_{n-p}^2=0.$$

同理可证 2）当 $s<0$ 时，即 $1/2(n-|s|)=p$ 的情形.

4.5.3 正定二次型

定义 4.5.4 设 $f(x_1,\cdots,x_n)$ 为 n 元实二次型. 对 n 个不全为零的实数 $c_1,\cdots,c_n$，

1）若都有 $f(c_1,\cdots,c_n)>0$，则称 $f(x_1,\cdots,x_n)$ 为**正定二次型**；

2）若都有 $f(c_1,\cdots,c_n)<0$，则称 $f(x_1,\cdots,x_n)$ 为**负定二次型**；

3）若都有 $f(c_1,\cdots,c_n)\geqslant 0$，则称 $f(x_1,\cdots,x_n)$ 为**半正定二次型**；

4）若都有 $f(c_1,\cdots,c_n)\leqslant 0$，则称 $f(x_1,\cdots,x_n)$ 为**半负定二次型**；

5）若 $f(x_1,\cdots,x_n)$ 既不是半正定，又不是半负定的，则称 $f(x_1,\cdots,x_n)$ 为不定

二次型.

定理 4.5.5 设 $q(x_1,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=X'AX$,其中 $X'=(x_1,\cdots,x_n)$,$A=(a_{ij})=A'$,则总可通过坐标的正交变换 $\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix}=U\begin{pmatrix}y_1\\ \vdots\\ y_n\end{pmatrix}$ 化为

$$\lambda_1y_1^2+\lambda_2y_2^2+\cdots+\lambda_ny_n^2 \tag{4.1}$$

其中 U 为 n 阶正交阵,而 $\lambda_i\in\mathbf{R},i=1,\cdots,n$,是 A 的全部特征值,这就是二次型的化主轴问题,即任一实二次型均可通过正交的变量替换化为标准形(4.1).由此即得:

1) 二次型的秩是它的矩阵 A 的非零特征值的个数;

2) 二次型的符号差是它的矩阵 A 的正特征值的个数与负特征值的个数的差.

例 4.5.3 设 A 为 n 阶正定阵,B 为 n 阶半正定阵.试证:$|A+B|\geqslant|A|$.

证明 由 A 是 n 阶正定阵知 A 合同于 n 阶单位阵 I_n,即存在 n 阶实可逆阵 P,使 $P'AP=I_n$.又由 B 为 n 阶半正定阵知 $P'BP$ 仍是半正定阵.所以存在 n 阶正交阵 Q,使得 $Q'(P'BP)Q=\mathrm{diag}(\lambda_1,\lambda_2,\cdots,\lambda_n)$,其中 $\lambda_i(i=1,2,\cdots,n)$ 是 $P'BP$ 的全部特征根,因而 $\lambda_i\geqslant 0,i=1,2,\cdots,n$. $T=PQ$,于是

$$T'AT=Q'(P'BP)Q=Q'Q=I_n,$$

$$T'(A+B)T=T'AT+T'BT=\mathrm{diag}(1+\lambda_1,\cdots,1+\lambda_n).$$

由此即得 $|A+B|=1/|T|^2(1+\lambda_1)(1+\lambda_2)\cdots(1+\lambda_n)\geqslant 1/|T|^2=|A|$.

4.5.4 正定矩阵及应用

定理 4.5.6 设 A 为 n 阶正定矩阵,则

1) A^{-1} 是正定阵;

2) $kA(k>0)$ 是正定阵;

3) A^* 是正定阵;

4) A^m(m 为任意整数)是正定阵;

5) 若 B 亦是 n 阶正定阵,那么 $A+B$ 也是正定阵.

定理 4.5.7 设 n 元实二次型 $q(x_1,\cdots,x_n)=\sum_{i=1}^{n}\sum_{j=1}^{n}a_{ij}x_ix_j=X'AX$,其中 $X'=(x_1,\cdots,x_n)$,$A=(a_{ij})=A'$ 是正定的,则下述诸命题彼此等价:

1) 它的秩与符号差都等于 n;

2) 它的矩阵 A 合同于 n 阶单位阵 I_n,即存在 n 阶可逆阵 P,使 $P'AP=I_n$;

3) 它的矩阵 A 的一切顺序主子式都大于零;

4）它的矩阵 $A = Q'Q$，其中 Q 为 n 阶实可逆阵；

5）它的矩阵 A 的特征值都是正数；

6）它的矩阵 $A = S^2$，其中 S 为正定阵.

实二次型 $q(x_1, \cdots, x_n) = X'AX$ 是半正定的，则有平行的结果.

定理 4.5.8　设 A 是 n 阶正定矩阵，a 为实数，B 是非零实数列向量，设线性方程组 $(A + aI)X = B$ 的解为 $X = X(a)$，则 $\varphi(a) = |X(a)|$ 是$[0, +\infty)$上的严格递减函数.

分析　首先确定 $\varphi(a)$ 是用含参数 a 的向量的长 $|X(a)|$ 来定义的，而 $X(a)$ 是线性方程组$(A+aI)X = B$的解. 由 A 正定知存在正交阵 U，使得 $U'AU$ 为对角形，且对角线上的元素都是正实数，应该由此入手来讨论.

证明　因为 A 是正定矩阵，存在正交阵 U，使得 $U'AU = \mathrm{diag}(\lambda_1, \lambda_2 \cdots \lambda_n)$（表示对角线上元素为 $\lambda_1, \lambda_2, \cdots, \lambda_n$ 的对角形矩阵，且 $\lambda_1, \lambda_2 \cdots \lambda_n$ 为正实数）. 用

$$A = U\mathrm{diag}(\lambda_1, \lambda_2, \cdots, \lambda_n)U',$$

代入线性方程组：$(A + aI)X = B$，即

$$U\mathrm{diag}(\lambda_1 + a, \lambda_2 + a, \cdots, \lambda_n + a)U'X = B,$$

则 $\mathrm{diag}(\lambda_1 + a, \lambda_2 + a, \cdots, \lambda_n + a)U'X = U'B$. 作变量代换 $Y = U'X$，令 $C = U'B$，则 $|Y| = |X|$. 设 $C = (c_1, c_2 \cdots c_n)'$，$Y = (y_1, y_2 \cdots y_n)'$，则

$$\mathrm{diag}(\lambda_1 + a, \lambda_2 + a, \cdots, \lambda_n + a)Y = C.$$

所以 $y_i = c_i'(\lambda_i + a)$，其中 $\lambda_i > 0$，$a \geqslant 0$，所以 $\lambda_i + a > 0$，则

$$\varphi(a) = |X(a)| = |Y(a)| = \left[\sum_{i=1}^{n}\left(\frac{c_i}{\lambda_i + a}\right)^2\right]^{1/2},$$

当 $c_i \neq 0$ 时，$\left|\dfrac{c_i}{\lambda_i + a}\right|$随 a 的增大而减小；

当 $c_i = 0$ 时，这一项为零，又由 $C = U'B$ 知，当 b_i 不全为零时，c_i 一定不全为零，所以 $\sum\limits_{i=1}^{n}\left(\dfrac{c_i}{\lambda_i + a}\right)^2$ 随 a 的增大而减少，则 $\varphi(a) = |X(a)|$ 是$[0, +\infty)$上的严格递减函数.

定理 4.5.9　设 $q(X) = X'AX$，其中 $A = (a_{ij})$ 为 n 阶实对称阵，$X = (x_1, \cdots, x_n)'$，则

1）如果 $a_{ij} = \dfrac{1}{i + j}$，则 A 为正定矩阵；

2）如果 A 为正定矩阵，则 $B = \left(\dfrac{a_{ij}}{i + j}\right)_{n\times n}$ 也是正定矩阵.

证明　1）$q(X) = \sum\limits_{i,j=1}^{n} a_{ij}x_ix_j = \sum\limits_{i,j=1}^{n}\dfrac{x_ix_j}{i + j} = \sum\limits_{i,j=1}^{n} x_ix_j\int_0^{\infty} \mathrm{e}^{-(i+j)t}\mathrm{d}t$

$$= \int_0^{\infty} \left(\sum_1^n x_i e^{-it} x_j e^{-jt} \right) dt = \int_0^{\infty} \left(\sum_{k=1}^n x_k e^{-kt} \right)^2 dt,$$

对于一切 $X \neq 0$,被积函数都大于零,所以积分值大于零,即对于任何不为零的 X,都使 $q(X) = X'AX > 0$,则 $q(X)$ 为正定二次型,A 为正定阵.

2) 设 $q(X) = X'BX = \frac{1}{i+j} X'AX = \sum_{i,j=1}^n a_{ij} x_i x_j \int_0^{\infty} e^{-(i+j)t} dt = \int_0^{\infty} \left(\sum_{i,j=1}^n a_{ij} x_i x_j e^{-(i+j)t} \right) dt$,

令 $Y = (x_1 e^{-t},\ x_2 e^{-2t} \cdots\ x_n e^{-nt})$,则上面广义积分中被积函数为 $Y'AY$,对于任意不为零的 X,Y 也不为零,而 A 为正定矩阵,故二次型 $Y'AY$ 正定,即对于 $\forall Y$ 不为零,$Y'AY > 0$ 成立.

综上所述,$\forall X \neq 0$,都有 $Y \neq 0$,使得 $Y'AY > 0$,即被积函数总大于零,所以积分值一定为正值,所以 $q(X) = X'BX$ 为正定二次型,B 为正定矩阵.

例 4.5.4　已知 A、C 为 n 阶实对称阵,且 C 为正定矩阵,若矩阵方程 $AX + XA = -C$ 有唯一解 B,且 A 的特征值为互不相等的负数,求证: B 为正定矩阵.

证明　先证 B 是实对称阵,由 $AB + BA = -C$ 知 $B'A' + A'B' = -C$,即 $B'A + AB' = -C$,故 B' 也为矩阵方程的解,由解的唯一性知: $B' = B$,即 B 对称.

再证 B 是正定矩阵,因为 A 是实对称阵,一定存在正交阵 U,使得

$$U'AU = \text{diag}(\lambda_1, \lambda_2 \cdots \lambda_n),\text{其中 } \lambda_i < 0.$$

令 $U'BU = (b_{ij})_{n\times n}$, $U'CU = (c_{ij})_{n\times n}$, 则 $U'AU \cdot U'BU + U'BU \cdot U'AU = -U'CU$,即

$$\text{diag}(\lambda_1, \lambda_2 \cdots \lambda_n)(b_{ij})_{n\times n} + (b_{ij})_{n\times n} \text{diag}(\lambda_1, \lambda_2 \cdots \lambda_n) = -(c_{ij})_{n\times n}.$$

经计算可得 $c_{ij} = \lambda_i b_{ij} + \lambda_j b_{ji}$,又 $\lambda_i + \lambda_j < 0$,所以 $b_{ij} = \frac{-c_{ij}}{\lambda_i + \lambda_j}$,即

$$U'BU = \left(\frac{-c_{ij}}{\lambda_i + \lambda_j} \right),\ \forall X \neq 0.$$

二次型 $q(X) = X'(b_{ij})_{n\times n} X$

$$= \sum_{i,j=1}^n x_i x_j \left(\frac{-c_{ij}}{\lambda_i + \lambda_j} \right) = \sum_{i,j=1}^n c_{ij} x_i x_j \int_0^{\infty} e^{(\lambda_i + \lambda_j)t} dt = \int_0^{\infty} \sum_{i,j=1}^n c_{ij} (x_i e^{\lambda_i t})(x_j e^{\lambda_j t}) dt.$$

令

$$Y = (x_1 e^{\lambda_1 t}, x_2 e^{\lambda_2 t} \cdots x_n e^{\lambda_n t})',\ q(X) = \int_0^{\infty} Y'(c_{ij}) Y dt > 0.\ (\text{当 } X \neq 0 \text{ 时}, Y \neq 0)$$

则 $(b_{ij})_{n\times n}$ 为正定矩阵,从而由 $(b_{ij})_{n\times n} = U'BU$ 知 B 为正定矩阵.

第 5 章 群论基础

群论是现代数学中最早和内容最丰富的分支之一.关于群的理论的系统研究早在19世纪初就已开始,其中,变换群在几何学中起着重要的作用,而有限群是方程论中的Galois理论的基础,这两个领域对群论的系统研究提供了原动力.

5.1 群论基础

5.1.1 群、子群、同态

定义 5.1.1 群是一个非空集合 G,加之一个二元运算"$\cdot$".它满足以下性质:

1)(结合性) 对于 $\forall a, b, c \in G, (a \cdot b) \cdot c = a \cdot (b \cdot c)$.

2)(单位元) 对于 $\forall a \in G, \exists e \in G$,使得 $e \cdot a = a \cdot e = a$.

3)(逆元) 对于 $\forall a \in G, \exists a^{-1} \in G$,使得 $a \cdot a^{-1} = a^{-1} \cdot a = e$.

如果群 G 的运算满足交换律,则称该群为**交换群**或 **Abel 群**.

Q,**R**,**C** 对于加法均是 Abel 群.

设 G 是群,若集合 G 有限,称 G 为**有限群**,否则叫**无限群**.若有限群 G 共有 n 个元素,则称 G 为 n 阶群或 n 元群,称 $n = |G|$ 为有限群 G 的**阶**.

例 5.1.1 **Q** 上 n 阶非退化矩阵的全体 $GL_n(Q)$ 对于矩阵的乘法作成群,**R** 上 n 阶正交矩阵 O_n 的全体,对于矩阵的乘法作成群.这些矩阵群不是交换群.

例 5.1.2 欧氏平面中保持欧氏距离不变的运动叫做欧氏运动.由于欧氏运动必为到自身的一一对应,并且它的逆仍是欧氏运动,而两个欧氏运动的合成仍是欧氏运动,从而全体欧氏运动形成群,叫做平面上的欧氏运动群,这也是非Abel群.

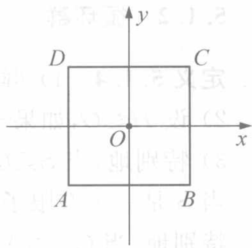


图 5-1

平面上点绕坐标原点的旋转变换的集关于映射的积作成群,叫做平面上的旋转群.

$ABCD$ 是一个固定的正方形(图 5-1),令 R 为该正方形绕中心 O 逆时针方向旋转 $90°$ 的变换,则 $R^4 = RRRR = E$ 是恒

等变换. 令 H 为正方形 $ABCD$ 对 x 轴的反射，令 V 为正方形 $ABCD$ 对 y 轴的反射，令 D 为正方形 $ABCD$ 对 BD 轴的反射，D' 为对 AC 轴的反射，则 $D_4=\{R_1,R_2,R_3,R_4=E,H,V,D,D'\}$ 关于变换的积是一个群. $R_4=E$ 是它的恒等元. D_4 叫做二面体群.

定义 5.1.2 设 $(G,\cdot)$ 是一个群，G 的非空子集 A 关于二元运算"$\cdot$"封闭，且 $(A,\cdot)$ 也是一个群，则称 $(A,\cdot)$ 为群 $(G,\cdot)$ 的**子群**，并记为 $A\leqslant G$，则 $n\mathbf{Z}\leqslant\mathbf{Z}$，$n\mathbf{Z}$ 与 $\mathbf{Z}$ 具有相同的单位元 0.

例 5.1.3 设 $(G,\cdot)$ 是一个群，令 $C=\{a|a\in G,x\in G\Rightarrow xa=ax\}$，则 $C\leqslant G$，C 叫 G 的中心.

定理 5.1.1 设 $(G,\cdot)$ 是一个群，e 是 G 的单位元，则 G 的子集 H 是子群的充要条件是：

(1) $a,b\in H\Rightarrow ab\in H$；

(2) $a\in H\Rightarrow a^{-1}\in H$.

从这个定理立即得到，若 H_1,H_2 是 G 的子群，则 $H_1\cap H_2$ 也是 G 的子群. 一般地，若 $\{H_i\mid i\in I\}$ 是 G 的子群族，则其交 $\bigcap\{H_i\mid i\in I\}$ 也是 G 的子群.

定义 5.1.3 设有群 $(G,\cdot)$，$(G',\odot)$，映射 $\mu: G\to G'$ 叫**同态**，是指对 $\forall a,b\in G$ 有 $\mu(a\cdot b)=\mu(a)\odot\mu(b)$.

从同态的定义立即得到同态的下述基本性质：

(1) 若 $\mu: G\to G'$ 是群同态，则 $\mu(e)=e'$，这里 e,e' 分别是 G,G' 的恒等元.

(2) 若 $\mu: G\to G'$ 是群的同态，$a\in G$，则 $\mu(a^{-1})=\mu(a)^{-1}$.

(3) 若 $\mu: G\to G'$ 是群的同态，$a\in G,k\in\mathbf{Z}$，则 $\mu(a^k)=\mu(a)^k$.

(4) 若 $\mu: G\to G'$ 是群的同态，则 $G_\mu=\{\mu(a)\mid a\in G\}$ 是 G' 的子群，叫做 G 关于 μ 的同态像.

定理 5.1.2(Cayley 定理) 任一群必同构于一个变换群.

证明 定义映射 $\rho: G\to \mathrm{Sym}(G)$，对 $a\in G$，$a\rho=\rho_a$，其中 $\rho_a: G\to G$，对 $x\in G$，$x\rho_a=xa$. 由于 ρ 为单同态，且 $G\cong G_\rho\leqslant\mathrm{Sym}(G)$. 而 G_ρ 是集合 G 上的变换群.

5.1.2 循环群

定义 5.1.4 1) 群 G 中含 S 的最小子群称为 S 在 G 中**生成的子群**，记作 $\langle S\rangle$；

2) 设 $H\leqslant G$，如果子集 $S\subseteq H$ 且 $\langle S\rangle=H$，则称 S 是子群 H 的一个**生成元集**；

3) 特别地，当 $S\subseteq G$ 而 $\langle S\rangle=G$ 时，称 S 生成群 G，而 S 是群 G 的一个生成元集.

当 S 是一个有限子集时，我们称 G 为有限生成的.

特别地，当 $G=\langle a\rangle$ 是由一个元素 a 组成时，称 G 是由 a 生成的**循环群**.

如果 G 是一个循环群，则 G 必有下列性质：

1) $G=\{\cdots,a^{-n},\cdots,a^{-2},a^{-1},e,a,a^2,\cdots,a^m,\cdots\}=\{a^n\mid n\in\mathbf{Z}\}$，其中 $a^n=a^m\Leftrightarrow n=m$；

2) $G=\{e,a,a^2,\cdots,a^{n-1}\}$,其中 $a^n=e$, 而 $a^s=a^t,0\leqslant s,t\leqslant n-1\Leftrightarrow s=t$.

例 5.1.4　$(\mathbf{Z},+)$是无限循环群,1,−1 都是生成元.$(\mathbf{Z}_n,+)$是一个 n 阶循环群.

令 U_n 是 1 的 n 次单位原根的乘法群,

令 $\theta=\dfrac{2\pi}{n}$,则 $U_n=\{e^{\frac{2k\pi}{n}i}\mid k=0,1,\cdots,n-1\}=\langle e^{i\theta}\rangle$ 是一个循环群.

循环群有且仅有$(\mathbf{Z},+)$ 和 $U_n,n\in\mathbf{N}$.

循环群是一类最简单的群.对循环群来说,很多关于群的基本问题都比较容易解决,例如,确定一个给定群的所有子群,这对一般群来说是一个很难解决的问题,但对于循环群来说却很容易解决.循环群的子群都是循环群.

1) 若 G 是无限循环群,则对每个正整数 m,G 恰有一个指数为 m 的子群 $G_m=<a^m>$,并且它们是 g 的全部子群.

2) 若 G 是 n 阶有限循环群,则对 n 的每个正因子 m,G 恰有一个指数为 m 的 n/m 阶子群 $G_m=<a^m>$,并且它们是 G 的全部子群.

5.1.3　轨道、子群的陪集

定义 5.1.5　设 G 是集 S 上的一个变换群,即 $\mathrm{Sym}(S)$ 的子群,则 G 确定了 S 中的一个二元关系:对于 $x,y\in S,x\sim y\Leftrightarrow$ 有 $\alpha\in G$,使得 $y=x\alpha$.这是一个等价关系,S 中元素 x 所在的等价类$[x]=xG=\{x\alpha\mid\alpha\in G\}$ 为 x 的 G 轨道,或简称**轨道**,于是集合 $\sum$ 分拆成一些轨道,在同一轨道中,可以通过某个 $\alpha\in G$ 的作用将其一个元素变为另一个元素,而不同轨道中的两个元素不可以这样做.

如果 G 在 S 上的作用只有一个轨道,则称 G 在 S 上是**传递**的.

例 5.1.5　设 $\alpha\in S_n$,且 $\alpha=(i_1i_2\cdots i_r)(j_1j_2\cdots j_b)\cdots(l_1l_2\cdots l_u)$ 是 α 分解为两两不相交的循环的乘积的分解式,且包括一切可能的 1 阶循环在内.令 $G=<\alpha>$ 是由 α 生成的循环群,则 $S=\{1,2,\cdots,n\}$ 关于 $G=<\alpha>$ 确定的 G 轨道正是$\{i_1i_2\cdots i_r\}$,$\{j_1j_2\cdots j_b\}$,$\cdots$,$\{l_1l_2\cdots l_u\}$.

例 5.1.6　设 G 是任一群,则由 Cayley 定理的证明知道,右平移的集 $G_\rho=\{\rho_a\mid a\in G\}$ 是 G 的一个变换群.由于对任何 $y\in G$,方程 $y=xa$ 有解,故 G_ρ 在 G 上是传递的.同样地,$G_\lambda=\{\lambda_a\mid a\in G\}$ 在 G 上也是传递的.

定义 5.1.6　设 $H\leqslant G$,令 $H_\rho=\{\rho_a\mid a\in H\}$,则 $H_\rho\leqslant G_\rho$,从而 H_ρ 是 G 的一个变换群.设 $x\in G$,则它的 H_ρ 轨道正是 $xH=\{xh\mid h\in H\}$,它叫做 x 关于子群 H 的**左陪集**,并称 G 的分类$\{xH\mid x\in H\}$ 为 G 关于 H 的**左陪集分解**,常记作 $L(H)$.

例 5.1.7　考虑 $(\mathbf{Z},+)$,对 $n\in\mathbf{N}$,令 $H=n\mathbf{Z}=\{nk\mid k=0,\pm1,\pm2,\cdots\}$,则 a,$b\in\mathbf{Z}$ 在关于 $H=n\mathbf{Z}$ 的同一左陪集中的充要条件是 $n\mid(a-b)$,即 $a\equiv b(\mathrm{mod}\,n)$,从而 $\mathbf{Z}$ 关于 $n\mathbf{Z}$ 的左陪集分解正是 $\mathbf{Z}$ 关于模 n 同余的剩余类的集 $\mathbf{Z}/<n>=\{[0],[1],\cdots,$

$[n-1]\}$.

定理 5.1.3(Lagrange 定理)　设 G 为有限群,H 是它的子群,则 $|G|=|H|[G:H]$.

证明　设 $|G|=n$, $|H|=m$,则对 $\forall x\in G$, $|xH|=m$,设 $\{x_1H, x_2H, \cdots, x_rH\}$ 是 G 关于 H 的左陪集分解,于是 $n=mr$,此处 $r=[G:H]$.

推论 5.1.4　若 $|G|=n$,则 G 的任一子群的阶是 n 的一个正因子,$x\in G$,则 $x^n=e$.

Lagrange 定理是群论中简单而有用的定理,例如,一个 6 阶群只能有 1,2,3,6 阶子群,不能有 4 阶或 5 阶子群;素数阶群只有平凡子群.

如果在上面的讨论中考虑左平移的群 $G_\lambda=\{\lambda_a \mid a\in G\}$ 以及子群 $H_\lambda=\{\lambda_a \mid a\in H\}$,则得到 $x\in G$ 关于 H_λ 的轨道 $Hx=\{hx \mid h\in H\}$,它叫做 x 关于 H 的右陪集. 与左陪集的情形一样,每一右陪集与 $He=H$ 有相同的基数,从而每一左陪集与每一右陪集有相同的基数.

另外,考虑左陪集分解 $L(H)=\{xH \mid x\in G\}$ 到右陪集分解及 $R(H)=G_\lambda=\{Hx \mid x\in G\}$ 的映射 μ,对 $xH\in L(H)$, $\mu(xH)=Hx^{-1}$,则 μ 是一个双射,从而 $L(H)$ 与 $R(H)$ 有相同的基数,于是符号 $[G:H]$ 也适用于右陪集分解.

5.1.4　正规子群、商群和同态定理

若 η 是半群 S 上的一个同余,则商集 S/η 关于运算 $\eta(a)\eta(b)=\eta(ab)$ 是一个半群,叫做 S 关于 η 的**商半群**.

下面我们讨论当半群 S 是一个群 G 的情形,首先我们有下述定理:

定理 5.1.5　若 η 是群 G 上的一个同余,则 G/η 是一个群.

证明　由上面的讨论我们知道,G/η 是一个半群,设 e 是群 G 的恒等元,则对任何 $\eta(a)\in G/\eta$, $\eta(a)\eta(e)=\eta(ae)=\eta(a)$, $\eta(e)\eta(a)=\eta(ea)=\eta(a)$,而 $\eta(e)$ 是 G/η 的恒等元. 又设 $\eta(a)\in G/\eta$,则 $\eta(a)\eta(a^{-1})=\eta(aa^{-1})=\eta(e)$, $(a^{-1})(a)=(a^{-1}a)=e$,从而 $a^{-1}\eta$ 是 $a\eta$ 的逆元,即 $(a\eta)^{-1}=a^{-1}\eta$,从而 G/η 是一个群. 常称 G/η 为群 G 关于同余 η 的商群.

例 5.1.8　在 $(\mathbf{Z},+)$ 中,"模 n 同余",即 $a\equiv b(\mathrm{mod} n)$,是一个同余,从而 $(\mathbf{Z}/\langle n\rangle,+)$ 是 $(\mathbf{Z},+)$ 关于这个同余的商群.

今设 η 是群 G 上的一个同余,设 $[e]=\{a \mid a\in G, a\eta e\}$,这里 e 是群 G 恒等元,则 $[e]$ 是 G 的一个子集,且 $e\in[e]$,若 $a,b\in[e]$,则 $a\eta e$, $b\eta e$,故 $ab\eta ee$,从而 $ab\in[e]$,又若 $a\in[e]$,则 $\eta(a)=\eta(e)$,故 $\eta(a^{-1})=\eta(a)^{-1}=\eta(e)^{-1}=\eta(e)$,从而 $a^{-1}\in[e]$,这样 $K=[e]$ 是 G 的一个子群.

另外,若 $a\eta b$,则 $\eta(a)=\eta(b)$,从而 $\eta(a)^{-1}\eta(b)=\eta(e)$,但

$$\eta(a)^{-1}\eta(b)=\eta(a^{-1})\eta(b)=\eta(a^{-1}b),$$

从而 $a^{-1}b \in K = \eta(e)$，因此存在 $k \in K$ 使 $b = ak$，这样，同余关系 η 正是 K_η 等价，而$[a]$正是 a 关于子群 K 的左陪集，从而$[a]$即是 a 所在的右陪集 Ka，于是 $aK = [a] = Ka$，也就是说 $K = [e]$ 是下述定义给出的正规子群.

定义 5.1.7　设 $H \leqslant G$，对 $\forall a \in G$，有 $aH = Ha$，称 H 为 G 的**正规子群**，记作 $H \lhd G$.

现在我们可以给出群 G 中的同余关系与 G 的正规子群间的一个内在联系.

设 G 是一个群，η 是 G 中的一个同余关系，则等价类 $K = [e]$ 是 G 的一个正规子群，且对 $\forall a \in G$，$a\eta = aK = Ka$；反之，若 K 是 G 的一个正规子群，今定义 G 中的关系 η 为 K_η 等价，则 η 是一个同余关系，且由这个同余关系确定的等价类 $a\eta = aK = Ka$.

定理 5.1.6　对 G 的子群 H，$H \lhd G \Leftrightarrow$ 对 $\forall a \in G$，有 $a^{-1}Ha = \{a^{-1}ha \mid h \in H\} \subseteq H$.

定理 5.1.7　$f: G \sim H$，则：1) $\mathrm{Im}f$ 是群 H 的子群；2) $\mathrm{Ker}f$ 是群 G 的正规子群.

定理 5.1.8(群的第一同态定理)

1) 若 $H \lhd G$，则 $\varphi: G \to G/H$，$g \mapsto gH$ 是群 G 到其商群 G/H 的满同态；

2) 设 f 是群 G 到群 $\overline{G} = \mathrm{Im}f$ 的满同态，而 $H = \mathrm{Ker}f$，则 $G/H \cong \overline{G}$.

定理 5.1.9(群的第二同态定理)　设 f 是群 G 到群 $\overline{G}$ 上的满同态，$H = \mathrm{Ker}f$. 令 $L(G,H) = \{G$中所有包含 H 的子群$\}$，$L(\overline{G}) = \{\overline{G}$中所有子群$\}$，则 $\theta: L(G, H) \to L(\overline{G})$.

1) $S \mapsto \varphi(S) = \{\varphi(s), s \in S\} = \overline{S}$ 是集 $L(G,H)$ 到集 $L(\overline{G})$ 上的一个一一对应，且有 $S \supseteq T$ 当且仅当 $f(S) \supseteq f(T)$；

2) S 是 G 的正规子群当且仅当 $f(S)$ 是 $\overline{G}$ 的正规子群；

3) 当 S 是 G 的正规子群时，有 $G/S \cong \overline{G}/f(S)$.

5.1.5　对称群与置换群

除了同态基本定理，研究群的另一个重要手段是群在集合上的作用，或者说是群的置换表示，即一个给定群到某个置换群上的同态. 为此，我们在本节介绍有限集合上置换群的基本知识.

定义 5.1.8　设 S 是非空集合，集合 S 到自身的每一个一一对应叫做 S 上的一个**置换**.

以 $A(S)$ 表示上面全部置换构成的集合. 在函数合成运算 $f \circ g$ 之下，$A(S)$ 是群，称为 S 上的**对称群**，它的每个子群均叫集合 S 上的**置换群**.

设 S 和 S' 是两个有限集合，如果 $|S| = |S'| = n$，易知 $A(S)$ 和 $A(S')$ 同构，从而可以谈 n 元集合上的对称群，表示成 S_n，且 $|S_n| = n!$. 而 S_n 的每个子群均叫 n 元集合上的置换群.

我们以$(i_1 i_2 \cdots i_r)$表示如下的置换：$i_1 \to i_2, i_2 \to i_3, \cdots, i_r \to i_1$

把$(i_1 i_2 \cdots i_r)$叫做一个长为r的**轮换**. 由此，每个置换均可写成一些轮换的乘积，使得不同的轮换中没有相同元素. 例如

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 3 & 6 & 7 & 1 & 2 & 4 & 8 \end{pmatrix} = (1\ 5)(2\ 3\ 6)(4\ 7)(8).$$

除次序不同外，置换表示成没有共同数字的轮换之积是唯一的. 长为 2 的轮换叫**对换**.

容易看到，每个轮换可表成一些对换之积.

例如，$(i_1 i_2 \cdots i_r) = (i_1\ i_2)(i_1\ i_3)\cdots(i_1\ i_r)$，所以每个置换总可表成有限个对换之积. 这种表达式（乃至对换的个数）显然不唯一，但是，同一个对换以多种方式表成对换之积时，其所含对换个数的奇偶性是不变的. 表成奇（偶）数个对换之积的置换叫做**奇(偶)置换**.

两个奇置换或偶置换之积是偶置换，一个奇置换与一个偶置换之积是奇置换.

全体偶置换构成的子群叫做n元集合上的交错群A_n.

易知A_n是S_n的正规子群，并且$[S_n : A_n] = 2$，$|A_n| = n!/2\ (n \geqslant 2)$.

定义 5.1.9 只有平凡正规子群的群叫做**单群**.

素数阶群$\mathbf{Z}_p$是循环群，它只有平凡子群$\{1\}$和$\mathbf{Z}_p$，从而是单群. 元素个数大于 1 的群是单群的充要条件是它为素数阶(循环)群.

定理 5.1.10 当$n \geqslant 5$时，交错群A_n是单群.

我们要给出的证明是非常初等的. 首先需要两个引理.

引理 5.1.11 当$n \geqslant 3$时，长为 3 的轮换形成A_n的一个生成元系.

证明 设$\sigma \neq 1$是偶置换，则σ是偶数个对换之积. 从而只需证任意两个对换之积可用长为 3 的轮换表示即可.

对于$\tau = (ij)(rs)(i \neq j, r \neq s)$，如果$(ij) = (rs)$，则$\tau = 1$；如果$j = r, i \neq s$，则$\tau = (jsi)$；如果$i, j, r, s$两两不等，则$\tau = (ris)(ijr)$.

引理 5.1.12 如果N是$A_n(n \geqslant 3)$的正规子群，并且N包含一个 3 -轮换，则$N = A_n$.

证明 如果$(rsc) \in N$，则对每个$k \neq r, s, c$，$(rsk) = (rs)$.

$$(ck)(rsc)^2(ck)(rs) = [(rs)(ck)](rsc)^2[(rs)(ck)]^{-1} \in N.$$

则$N = A_n$.

定理 5.1.10 的证明 设$\{1\} \neq N \lhd A_n$，我们证明$N = A_n$.

N中必包含一个元素是长为 3 的轮换. 事实上，设$1 \neq \sigma \in N$，并且σ将$\{a_1, \cdots, a_n\}$中尽可能多的元素保持不动. 设σ恰好变动 3 个a_i，从而必是长为 3 的轮换. 首先，σ至少

变动3个a_i(因为只变动两个a_i的为对换,而对换是奇置换,不属于A_n),现在把σ写成没有公共元素的轮换之积,并且把最长的轮换写在左边.若σ恰好变动4个a_i,则$\sigma=(a_1a_2)(a_3a_4)$.由于$n\geqslant 5$,从而$\beta=(a_3a_4a_5)\in A_n$,而$\sigma_1=\beta\sigma\beta^{-1}=(a_1a_2)(a_4a_5)\in N$,于是$\sigma\sigma_1=(a_3a_4)(a_4a_5)=(a_3a_4a_5)\in N$,即是长为3的轮换.若$\sigma$至少变动5个$a_i$,则又分三种情形考虑:

1) σ包含长度$\geqslant 4$的轮换,即$\sigma=(a_1a_2a_3a_4\cdots)\cdots$.取$\beta=(a_2a_3a_4)\in A_n$,则$\sigma_1=\beta\sigma\beta^{-1}=(a_1a_3a_4a_2\cdots)\cdots\in N$,而$i\geqslant 5$时,$\sigma(a_i)=\sigma_1(a_i)$,从而$N$中$\sigma_1\sigma^{-1}$至多变动4个$a_i$,这与$\sigma$变动$a_i$个数的极小性矛盾;

2) σ中轮换最大长度为3,则$\sigma=(a_1a_2a_3)(a_4a_5\cdots)\cdots$,由于$\sigma$至少变动5个$a_i$,从而$\sigma$不是长为3的轮换.这样的$\sigma$至少变动6个$a_i$.取$\beta=(a_2a_3a_4)\in A_n$,则$\sigma_1=\beta\sigma\beta^{-1}=(a_1a_3a_4)(a_2a_5\cdots)\cdots\in N$,而$N$中置换$\sigma_1\sigma^{-1}$至多变动5个$a_i$,导致矛盾;

3) 设σ是一些对换之积:$\sigma=(a_1a_2)(a_3a_4)\cdots$,它至少变动6个$a_i$.取$\beta=(a_2a_3a_4)\in A_n$,则$\sigma_1=\beta\sigma\beta^{-1}=(a_1a_2)(a_4a_5)\in N$,而$\sigma\sigma_1{}^{-1}(\in N)$只变动4个$a_i$,矛盾;

综上可知,N中包含元素是长为3的轮换.由引理即知$N=A_n$.

5.1.6　群在集合上的作用

我们说过同态是研究群之间关系的基本手段.为了研究一个群G,自然希望有一些理想的"样板"群作为标准,然后通过研究G到样板群的各种同态来把握G的特性.理想的样板群有两类,一类是置换群,另一类是矩阵群.一个群G到置换群的同态叫G的置换表示,而到矩阵群的同态叫线性表示.研究群的线性表示是群论的一个美妙的分支,即通常所谓群表示理论,它在物理、化学、力学等许多方面都得到了重要应用.下面介绍群的置换表示理论的一些基本知识.

定义 5.1.10　设$\sum$是一个集合,$S(\sum)$是$\sum$上的对称群,群G到$S(\sum)$的每个同态$f: G\to S(\sum)$都叫做群G在集合$\sum$上的一个**置换表示**.

如果f是单同态,则称f是**忠实表示**.

这时,对于G中不同的元素g,$f(g)$是$\sum$上不同的置换.群G借助于置换表示f作用在集合$\sum$之上,也就是说,元素$g\in G$在集合$\sum$上的作用看成是置换$f(g)$,对于每个$a\in\sum$,定义$ga=f(g)a$.

设$\pi: G\to S(\sum)$是一个置换表示.在$\sum$上定义如下的关系:对于$a,b\in\sum$,$a\sim b\Leftrightarrow$有$g\in G$,使得$ga=b$.这是一个等价关系,因为

(1) $\pi(1_G)$是$\sum$的恒等置换,从而对每个$a\in\sum$,$1_G\cdot a=a$,即$a\sim a$;

(2) 如果 $a \sim b$,则有 $g \in G$ 使得 $ga = b$,于是 $g^{-1}b = a$,即 $b \sim a$;

(3) 若 $a \sim b, b \sim c$,则 $ga = b, hb = c$,其中 $g, h \in G$,于是 $(hg)a = c$,所以 $a \sim c$.(简言之,$\sim$ 是等价关系,因为 G 是群)

定义 5.1.11　对于上述等价关系,$\sum$ 中元素 a 所在的等价类是 $[a] = Ga = \{ga \mid g \in G\}$. 每个等价类叫一个 **$G$-轨道**,或简称**轨道**.

于是 $\sum$ 分拆成一些轨道,在同一轨道中,可以通过某个 $g \in G$ 的作用将其一个元变为另一个元,而不同轨道中的两个元不可以这样做.

例 5.1.9　设 G 是群,取 $\sum = G$,对每个 g、$a \in G$,可作如下映射: $f: G \to S(G)$, $f(g)a = ga$. 也就是说,对于 $g \in G$, $f(g)$ 是集合 G 上如下的置换: 它将 G 的每个元素 a 变成 ga.(由群 G 上的消去律可知 $f(g)$ 是 G 上的置换),由于 $(f(g)f(g'))a = f(g)(f(g')a) = f(g)(g'a) = gga = f(gg)a$,从而 $f(g)f(g') = f(gg')$,即 $f: G \to S(G)$ 是群的同态.

定义 5.1.12　f 是群 G 在集合 G 上的一个置换表示,这叫做群 G 的**左正则表示**. 由于 $g \in \mathrm{Ker}f \Leftrightarrow ga = a, \forall a \in G \Leftrightarrow g = 1_G$,于是 f 为单同态,即左正则表示是忠实的.

类似定义 $\tau: G \to S(G)$, $\tau(g) = ag^{-1}$, $\tau(g)\tau(g')a = \tau(g)a(g')^{-1} = a(g')^{-1}g^{-1} = a(gg')^{-1} = \tau(gg')a$,可知 τ 也是一个群同态. 表示 τ 叫做 G 的右正则表示,它也是忠实的.

定理 5.1.13[凯莱(Cayley)定理]　每个群均同构于某个置换群.

证明　由于正则表示 $f: G \to S(G)$ 是忠实的,根据同态基本定理,G 同构于 $f(G)$,而 $f(G)$ 是集合 G 上对称群 $S(G)$ 的子群,从而 $f(G)$ 是集合 G 上的置换群.

这个定理显示出置换群有资格作为一切群的样板群. 我们希望能给出群 G 在较小集合 $\sum$ 上的置换表示. 因为 $n = \left|\sum\right|$ 愈小,$S(\sum) = S_n$ 的子群愈容易研究.

例 5.1.10　设 $H \leqslant G$. 取 $\sum = \{aH \mid a \in G\}$,即 $\sum$ 是由 G 对于 H 的全部陪集 aH 构成的集合,定义 $f_H: G \to S(\sum)$, $f_H(g)(Ah) = gaH$, 即对每个 $g \in G$, $j_H(g)$ 把陪集 aH 变成 gaH,这是集合 $\sum$ 上的置换,并且 f_H 是群的同态,从而 f_H 给出的一个 G 置换表示,叫做群 G 对于子群 H 的左诱导表示. 对 $\forall a \in G$, 由于 $g \in \mathrm{Ker}f_H \Leftrightarrow gaH = aH$, $\forall a \in G \Leftrightarrow a^{-1}ga \in H \Leftrightarrow g \in aHa^{-1}$.

$\forall a \in G \Leftrightarrow g \in \bigcap\limits_{a \in G} aHa^{-1}$,从而 $\mathrm{Ker}f_H = \bigcap\limits_{a \in G} aHa^{-1} = H$ 的所有共轭子群的交.

类似可定义 G 对于子群 H 的右诱导表示:

$$\sum = \{Ha \mid a \in G\}, \tau_H: G \to S(\sum),\ \tau_H(g)(Ha) = Hag^{-1}.$$

$\mathrm{Ker}\tau_H$ 也是 H 的所有共轭子群的交.

例 5.1.11　设 A 是群 G 的任意子集,取 $\sum = \{aAa^{-1} \mid a \in G\}$(即 A 的全部共轭子集),定义 π: $G \to S(\sum)$,$\pi(g)(aAa^{-1}) = gaAa^{-1}g^{-1} = (ga)A(ga)^{-1}$,这是一个置换表示,叫做群 G 对于子集 A 的共轭表示. 由于 $g \in \mathrm{Ker}\pi \Leftrightarrow gaAa^{-1}g^{-1} = aAa^{-1}$, $\forall a \in G \Leftrightarrow a^{-1}ga \in N_G(A)$, $\forall a \in G \Leftrightarrow g \in aN_G(A)a^{-1}$, $\forall a \in G$,从而

$$\mathrm{Ker}\pi = \bigcap_{a \in G} aN_G(A)a^{-1}$$

为正规化子 $N_G(A)$ 的所有共轭子群的交.

定义 5.1.13　设群 G 作用于集合 $\sum$ 之上,则对每个元素 $a \in \sum$,$G_a = \{g \in G \mid ga = a\}$ 是 G 的一个子群,叫做元素 a 的**固定子群**.

(轨道公式)设有限群 G 作用于集合 $\sum$ 上,$a \in \sum$,则 $|G| = |G_a||[a]|$.

最后我们用群在集合上的作用解决一些群论问题.

引理 5.1.14　设 G 是 $2n$ 阶群,2 不整除 n,则 G 必有指数为 2 的正规子群.

定理 5.1.15　G 为有限群,$|G| \geqslant 6$ 且 $|G| \equiv 2(\mathrm{mod}4)$,则 G 不是单群.

定理 5.1.16　设 G 为有限群,p 是 $|G|$ 的最小素因子,如果 $N \leqslant G$,$[G:N] = p$,则 $H \lhd G$.

5.2　有限群的结构

5.2.1　Sylow 定理

拉格朗日定理是:若有限群 G 的阶数是 n,则 G 的每个子群的阶都是 n 的因子. 反过来,对于 n 的每个因子 d,G 未必有 d 阶子群. 例如,我们已知 60 阶群 A_5 是单群,它没有 30 阶子群,因为这样的子群一定是正规的.

但是下一定理表明,对于 $|G|$ 的特殊因子 d,G 必有 d 阶子群. 在本定理以及以后许多结果的证明中,我们不断使用群在集合上的作用这一有效工具.

定理 5.2.1　设 $p^r \| G|$,其中 p 为素数,以 $N(n)$ 表示 G 中 n 阶子群的个数,则 $N(p^r) \equiv 1(\mathrm{mod}p)$. 特别地,如果 $p^r \| G|$,则 G 至少存在一个 p^r 阶子群.

证明　令 $|G| = p^r n$,以 $\sum$ 表示 G 的全部 p^r 元子集组成的集族,则 $|\sum| = C_{np^r}^{p^r}$. 考虑 G 在 $\sum$ 上的如下作用 f: $G \to S(\sum)$,$f(g)M = Mg^{-1}$, $\forall g \in G, M \in \sum$. $\sum$ 分拆成一些轨道 T_i 之并 $\sum = \cup_i T_i$,$|\sum| = \sum_i |T_i|$,$|T_i| = |G:A_i|$,其中 $A_i = g \in G \mid M_i g^{-1} = M_i \mid$ 是轨道 T_i 中任一元素 M_i 的固定子群,由于 $M_i A_i = M_i$,M_i 可以

分拆成 $M_i = \bigcup_{j=1}^{k_i} g_{ij}A_i, g_{ij} \in M_i, 1 \leqslant j \leqslant k_i, k_i = |M_i| / |A_i| = p^r / |A_i|$. 于是 $|A_i| = p^{r_i}$, $r_i \leqslant r$. 如果 $r_i < r$, 则 $|T_i| = [G : A_i] = np^{r-r_i} \equiv 0 (\mathrm{mod}\, pn)$.
如果 $r_i = r$, 则 $|T_i| = n$. 于是 $C_{np^r}^{p^r} = |\sum| = \sum_i |T_i| \equiv \sum_{|T_i|=n} |T_i| = n = n\sum_{|T_i|} 1 (\mathrm{mod}\, pn)$.

现在计算 $\sum\limits_{|T_i|=n} 1$, 即长为 n 的轨道 T_i 的个数, 注意到

$$|T_i| = n \Rightarrow |A_i| = p^r \Rightarrow k_i = 1 \Rightarrow M_i = g_i A_i,$$

于是 p^r 阶子群 $B_i = g_i A_i g_i^{-1}$ 与 $M_i = g_i A_i$ 在同一轨道 T_i 之中, 并且若 $X \in T_i$, 则 $Xg = M_i = B_i g_i, X = B_i g_i g^{-1}$, 所以轨道 T_i 中 n 个元素即是 G 对于 p^r 阶子群 B_i 的 n 个陪集. 这 n 个陪集中除 B_i 外其余陪集不包含 1, 从而不会是子群.

G 的每个 p^r 阶子群均恰好在一个长为 n 的轨道之中, 于是 $\sum\limits_{|T_i|=n} 1 = N(p^r)$, 从而 $C_{np^r}^{p^r} \equiv nN(p^r)(\mathrm{mod}\, pn)$. 这个同余式对任意 $p^r n$ 阶群 G 均成立.

特别地, 取 G 为 $p^r n$ 阶循环群, 则它只有一个 p^r 子群, 代入上式即知 $C_{np^r}^{p^r} \equiv n(\mathrm{mod}\, pn)$. 于是 $n \equiv nN(p^r)(\mathrm{mod}\, pn)$, 从而 $N(p^r) \equiv 1(\mathrm{mod}\, p)$.

定义 5.2.1　设 G 为 $p^r n$ 阶群, 其中 p 为素数, $r \geqslant 1$, p 不整除 n, 则 G 的每个 p^r 阶子群均叫做 G 的 **Sylow$p$-子群**.

定理 5.2.2(Sylow 定理)　设 G 为有限群, 则

1) 对 $|G|$ 的每个素因子 P, 均存在 G 的 Sylow$p$-子群;

2) G 的 Sylow$p$-子群彼此共轭;

3) G 的 Sylow$p$-子群的个数$\equiv 1(\mathrm{mod}\, p)$;

4) 设 P 为 G 的一个 Sylow$p$-子群, 则 G 的 Sylow$p$-子群的个数为$[G : N_G(P)]$.

证明　1)和 3)由定理 5.2.1 直接推出, 由 2)容易得到 4), 从而只需证 2). 令 $\sum$ 是 G 的所有 Sylow$p$-子群构成的集合, 将 G 共轭作用于其上. 令 Δ 是一个 G-轨道. 取 $P \in \sum$, 再将 P 共轭作用于Δ 上. Δ 分拆成一些 P-轨道, 每个 P-轨道的长度是 $|P| = p^r$ 的因子. 如 $P \in \Delta$, 并且 P' 自身组成一个 P-轨道, 即 $Xp'x^{-1} = P', \forall x \in P$, 则 $P \leqslant N_G(P')$, 从而 $PP' \leqslant G$, 但是 $|PP'| = |P\|P'| / |P \cap P'|$ 仍为 P 的幂, 且 $P \leqslant PP'$, 由于 P 和 P' 均是 Sylow$p$-子群, 故必 $P' = PP' = P$. 这就表明当 $P \in \Delta$ 时, Δ 中长为 1 的轨道只有$\{P\}$, 从而 $|\Delta| \equiv 1(\mathrm{mod}\, p)$. 当 $P \notin \Delta$ 时, Δ 没有长为 1 的轨道, 从而 $|\Delta| \equiv 0(\mathrm{mod}\, p)$. 这两种情形不可能同时发生, 所以只能是所有 Sylow$p$-子群均在 Δ 中, $\sum = \Delta$, 即 G 在 $\sum$ 上的共轭作用是传递的, 即 G 的所有 Sylow$p$-子群彼此共轭.

推论 5.2.3　设素数 p 是 $|G|$ 的因子，则群 G 的每个 p 方幂阶的子群 B 均包含在 G 的某个 Sylow$p$-子群内.

推论 5.2.4　设 P 是 G 的 Sylow$p$-子群，$A \leqslant G$，且 $N_G(P) \leqslant A$，则 $N_G(A) = A$.

推论 5.2.5[弗拉梯尼(Fratini)定理]　M 是 G 的正规子群，P 为 M 的 Sylow$p$-子群，则 $G = MN_G(P)$.

例 5.2.1　148 阶群不是单群.

证明　取 $p = 37 \mid 148$，则 $N(37) \equiv 1(\text{mod}37)$，从而 $N(37) = 37h + 1$，由于 148 阶群 G 的全部 Sylow$p$-子群形成一个共轭类，其总数应当是 $|G| = 148$ 的因子，即 $N(37) = 37h + 1 \mid 148$，于是 $37h + 1 \mid 4$，这只能 $h = 0$，即 $N(37) = 1$，因此 G 只有一个 37 阶子群，从而必然是正规子群，则 G 不是单群.

例 5.2.2　56 阶群 G 不是单群.

定理 5.2.6　设 p 和 q 是两个素数，则：

1) pq 阶群 G 不是单群；

2) p^2q 阶群 G 不是单群.

证明　1) 若 $p = q$，p^2 阶群 G 是 Abel 群，由定理 5.2.1 知它有 p 阶子群，Abel 群的子群都是正规的，所以 G 不是单群.

如果 $p \neq q$，不妨设 $p > q$，$N(p) = np + 1 \mid q$，而 $q < p$，只能 $n = 0$，G 只有一个 Sylow$p$-子群，它是正规子群，于是 G 不是单群.

2) 若 $p = q$，已证过 p^3 阶群 G 必有非平凡的中心 $C(G)$，且 $C(G)$ 有 p 阶子群 N，显然 N 是 G 的正规子群，因此 G 不是单群.

如果 $p > q$，则 $N(p^2) = np + 1 \mid q$，有 $n = 0$，G 有正规的 p^2 阶 Sylow 子群，G 不是单群. 如果 $p < q$，则 $N(q) = nq + 1 \mid p^2$，如果 $N(q) = 1$，则 G 有正规 q 阶子群，G 不是单群. 由于 $p < q$，所以 $N(q)$ 不能为 p. 最后若 $N(q) = p^2$，即 G 有 p^2 个 q 阶子群，它们共占据 G 的 $p^2(q-1)+1$ 个元素，余下 $p^2 - 1$ 个元素和 1_G 便构成 G 唯一的 p^2 阶 Sylow 子群 P，P 是 G 的正规子群，所以 G 也不是单群.

定理 5.2.7　非 Abel 单群的最小阶数是 60，且 60 阶单群必同构于 A_5.

5.2.2　有限 Abel 群的结构定理

本节我们将看到非常漂亮、完整的有限交换群的结构定理，由此我们将具体地理解什么是群的结构理论.

定义 5.2.2　设 G 是加群，而 $H_i(1 \leqslant i \leqslant s)$ 是 G 的子群，如果

1) $G = H_1 + \cdots + H_s$，即每一个 g 可表示成 $h_1 + \cdots + h_s$，$h_i \in H_i$；

2) 上面的表示方法是唯一的，即对 $\forall g \in G$，由 $g = h_1 + \cdots + h_s = h_1' + \cdots + h_s'$，其中 $h_i, h_i' \in H_i$，一定有 $h_i = h_i'$，$i = 1,2,\cdots,s$，则称 G 是子群 $H_1,\cdots,H_s$ 的(内)直和，

记为 $G = H_1 \oplus H_2 \oplus \cdots \oplus H_s$，也称 G 可以分解成 $H_1, \cdots, H_s$ 的直和.

引理 5.2.8　在加群 G 中，

1）若元素 g 的阶为 s，而 $(s,t) = 1$，则 tg 的阶亦为 s，且 $<g> = <tg>$；

2）若元素 g 的阶为 s，而 $(s,t) = d$，则 tg 的阶为 s/d；

3）若 $g_1 + \cdots + g_m = 0$，且 g_i 的阶 $s_i, i = 1,2,\cdots,m$，两两互素，则每个 $g_i = 0$.

定理 5.2.9　设 G 是有限加群，$|G| = n = p_1^{m_1} p_2^{m_2} \cdots p_t^{m_t}$，其中 p_i 是不同素数，则

1）$G = H_1 \oplus H_2 \oplus \cdots \oplus H_t$，其中 H_i 是 p_i-群，$i = 1,\cdots,t$；

2）若 $G = H_1 \oplus \cdots \oplus H_t = H_1' \oplus \cdots \oplus H_t'$，其中 H_i, H_i'是 p_i-群，$i = 1,\cdots,t$，则对 $\forall i$，有 $H_i = H_i'$. 设$\{g_1,\cdots,g_k\}$，$\{h_1,\cdots,h_k\}$是G的两个元素个数相等的生成元集，其相应的阶集依次为$\{p^{m_1},\cdots,p^{m_k}\}$，$\{p^{l_1},\cdots,p^{l_k}\}$，如果 $m_1 + \cdots + m_k < l_1 + \cdots + l_k$，我们就说生成元集$\{g_1,\cdots,g_k\}$ 小于生成元集$\{h_1,\cdots,h_k\}$.

定理 5.2.10　设 G 为有限 p-加群，$|G| = p^m$，则有

1）$G = \mathbf{Z}g_1 \oplus \cdots \oplus \mathbf{Z}g_k$；

2）若 $G = \mathbf{Z}g_1 \oplus \cdots \oplus \mathbf{Z}g_k = \mathbf{Z}h_1 \oplus \cdots \oplus \mathbf{Z}h_s$，则必有 $k = s$，且适当重排脚标后有 $\mathbf{Z}g_i \cong \mathbf{Z}h_i, i = 1,\cdots,k$.

定义 5.2.3(外直积的定义)　设群 $G_i, i = 1,\cdots,n$，令集合 $G = G_1 \times G_2 \times \cdots \times G_n$，而规定集 G 中的一个二元运算如下：对 $g_i, h_i \in G_i, i = 1,\cdots,n$，规定 $(g_1,\cdots,g_n) \cdot (h_1,\cdots,h_n) = (g_1h_1,\cdots,g_nh_n)$，这里 g_ih_i 是群 G_i 中的乘积.

直接验证 $(G, \cdot)$ 是一个群，称之为群 $G_1,\cdots,G_n$ 的（外）直积，记作 $G = G_1 \odot G_2 \odot \cdots \odot G_n$. 特别地，当所有 G_i 是交换群时，G 也是交换群. 这时我们常把 G 写成 $G = G_1 \oplus G_2 \oplus \cdots \oplus G_n$，称之为加群 G_i 的（外）直和，这时 G 的运算记作加法，写成 $(g_1,\cdots,g_n) + (h_1,\cdots,h_n) = (g_1 + h_1,\cdots,g_n + h_n)$，令 $G_i' = \{(0,\cdots,0,g_i,0,\cdots,0) \mid g_i \in G_i\}$，则 G_i 是 G 的子群，$G_i' \cong G_i$，且有 G 是其子群 $G_i'(i = 1,\cdots,n)$ 的（内）直和. 在这个意义上，内直和、外直和是互通的，虽然内直和概念是属于结构理论的，而外直和是属于构造理论的.

定理 5.2.11(有限交换群结构定理)　有限加群 G 可唯一地分解为素数幂循环群的直和，即设 $|G| = p_1^{m_1} p_2^{m_2} \cdots p_t^{m_t}$，$p_i$ 是不同素数，则

1）$G = G_{11} \oplus \cdots \oplus G_{1s_1} \oplus \cdots \oplus G_{t1} \oplus \cdots \oplus G_{ts_t}$，其中 G_{ij} 是阶为 $p_i^{m_{ij}}$ 的循环群.

2）自然数集$(p_1^{m_{11}},\cdots,p_1^{m_{1s_1}},p_t^{m_{t1}},\cdots,p_t^{m_{ts_t}})$ 由群 G 唯一确定.

这是一个很值得欣赏的结构定理，可以和算术基本定理相比，那里表示任意整数的是“素数”，构造方法是“乘积”，而这里表示任意有限加群的是“素数幂阶的循环群”，构造方法是“直和”.

自然数 n 的分解是 $n = p_1^{m_1} p_2^{m_2} \cdots p_t^{m_t}$，而有限加群 G 的阶 $|G| = n$ 的分解式为

$$|G| = n = p_1^{m_{11}} \cdots p_1^{m_{1s_1}} p_t^{m_{t1}} \cdots p_t^{m_{ts_t}}.$$

5.2.3　有限群的同构分类

代数学的基本问题之一就是同构分类问题.对于很多代数系统来说,这个问题已经解决.例如,线性代数中的向量空间,任意 n 维向量空间 $V_n(F)$ 都同构于 F^n,这实际上就是域上有限维向量空间的同构分类定理.它告诉我们,从同构的意义上来说,任意域上给定 n 维向量空间只有一个 F^n.

我们又有,给定正整数 n,任意两个 n 阶循环群彼此同构,即从同构的意义上说,n 阶循环群只有一个,那就是 $\mathbf{Z}_n$.这实际上就是有限阶循环群的分类定理.

基于同样的想法,Cayley 在给出了有限群的定义后,于 1878 年明确地提出了对于一般的 n 阶有限群的同构分类问题.与循环群的情形完全不同,人们发现这个问题是非常复杂和困难的.为了解决这个问题,许多数学家经过艰苦的努力,得到了若干个具有基本意义的有限群构造定理.根据这些定理,人们把解决有限群的同构分类问题归纳为两大步骤:① 分类所有的有限单群;② 找出"用单群来构造其他群"的方法.这就是所谓的 The Holder Program.它为解决有限群的同构分类问题指明了方向,并勾画了粗糙的轮廓.在某种意义上,有限单群是构造有限群的"基元素",为了使用这"基元素"来证明有限群的一般定理,我们需要了解关于有限单群的详尽信息.值得庆幸的是,分类所有的有限单群的工作,通过数百位数学家数十年的努力,已于 1980 年宣告完成.目前该项工作正在简化、整理.这一工作被认为是 20 世纪数学界获得的最伟大的成就之一.对于 The Holder Program 的第二部分,进展缓慢,仍有大量艰难工作要做.

定义 5.2.4　称群 G 的子群列 $1 = G_0 \leqslant G_1 \leqslant G_2 \leqslant \cdots \leqslant G_r = G$ 为 G 的**合成群列**.对 $\forall i$,每个商群 G_{i+1}/G_i 为单群.称上面的合成群列的长度为 r.合成群列中的商群称为这一群列的**合成因子**.

更一般地,一个群称为群 G 的合成因子,若它同构于 G 的某一合成列中的一个合成因子.

例 5.2.3　考虑群 S_5,它有一正规子群 A_5,且 A_5 是 S_5 的唯一的非平凡正规子群.由于 $S_5/A_5 \cong \mathbf{Z}_2$,也是单群,故 $1 \lhd A_5 \lhd S_5$ 为 S_5 的一个合成群列.

定义 5.2.5　称 N 为群 G 的**极大正规子群**,若 $N \lhd G$,且 G 没有真包含 N 的正规子群.

任意非平凡有限群均有极大正规子群.由对应定理易见,N 是 G 的极大正规子群当且仅当 G/N 是单群.正规的极大子群必然是极大正规子群(且指数为素数),但极大正规子群不一定是极大子群,因为它可能真含于一个非正规真子群中.

例 5.2.4　群 $1 \times \mathbf{Z}_2$ 是 $A_5 \times \mathbf{Z}_2$ 的极大正规子群,但不是极大子群.

定理 5.2.12　有限群有合成群列.

证明　设 G 是有限群,对 $|G|$ 作归纳法. 若 G 是单群,则 $1 \lhd G$ 是 G 的合成群列;否则 G 有极大正规子群 G',由归纳假设 G' 有合成群列 $1 = G_0 \leqslant G_1 \leqslant G_2 \leqslant \cdots \leqslant G_r = G$. 由于 G/G' 是单群,故 $1 = G_0 \leqslant G_1 \leqslant \cdots \leqslant G' \leqslant \cdots \leqslant G_r = G$ 是 G 的合成群列.

无限群不一定有合成群列. 例如,令 G 是无限循环群,因为 G 的每个非平凡子群均为无限循环群,因而均同构于 G. 由于 G 是非单群,故 G 没有子群是单群.

引理 5.2.13　设 $N \lhd G$,若 G 有合成群列,则 N 有合成群列.

定理 5.2.14(Jordan-Holder 定理)　若群 G 有合成群列,则 G 的任意两个合成群列有相同的长度,且它们等价.

定义 5.2.6　群 G 的一个子群列 $1 = G_r \leqslant \cdots \leqslant G_1 \leqslant G_0 \leqslant = G$ 称为**次正规群列**,若对 $\forall i$ 有 $G_{i+1} \lhd G_i$.

一个次正规群列称为**正规群列**,若对每个 i 有 $G_i \lhd G$.

合成群列是次正规群列的一个特例,但反之不然,因为次正规群列可能有平凡的商群因子,或者有非平凡商群因子,但不是单群.

若两个同长度的次正规群列满足合成群列等价定义中所给的条件,则称它们是等价的.

定义 5.2.7　一个给定的次正规群列经过插入子群所得到的新的次正规群列称为原来次正规群列的**加细**.

一个加细称为是真加细,若所插入的项中至少有一个在原来的群列中没有出现. 因而合成群列是没有重复项的且没有真加细的次正规群列.

定义 5.2.8　群 G 的正规群列称为**主群列**,若该正规群列中没有重复项,且任意两项之间不能再插入其他的正规子群.

定义 5.2.9　群 G 的主群列中相邻两群作成的商群称为 G 的**主因子**. 一个群称为群 G 的主因子,若它同构于群 G 的某一主群列中的一个主因子.

对主群列也有类似的 Jordan-Holder 定理,即一个群的任意两个主群列都有相同的长度,且等价. 其证明类似于合成群列情形的证明.

这两个结果是更一般结果的特殊情形,即算子群的 Jordan-Holder 定理.

定义 5.2.10　称 N 为群 G 的**极小正规子群**,若 $1 \neq N \lhd G$,且 N 不真包含 G 的非平凡正规子群.

任意非平凡有限群均有极小正规子群;单群有唯一的极小正规子群,即它自身.

定理 5.2.15　有限群有主群列.

证明　设 G 是有限群,对 $|G|$ 作归纳法,若 G 为单群,则 $1 \lhd G$ 是 G 的主群列,否则,G 有非平凡极小正规子群 N. 由归纳假设,G/N 有主群列,由对应定理可知这一主群列有形式:$1 = G_r/N \lhd \cdots \lhd G_1/N \lhd G_0/N = G/N$,其中对 $\forall i, G_i \lhd G$,且在 G_i 与 G_{i-1} 之间没有 G 的非平凡正规子群. 由于 N 是 G 的极小正规子群,故 $1 \lhd N = G_r \lhd \cdots$

$\lhd G_1 \lhd G_0 = G$ 是 G 的主群列.

由定义知,有限群的合成因子均是单群. 在本节的最后,我们来决定有限群的主因子. 首先有下面的引理

引理 5.2.16　G 有主群列,则群 G 的每个主因子是 G 的某个商群的极小正规子群.

证明　若 $1 = G_r \lhd \cdots \lhd G_1 \lhd G_0 = G$ 是 G 的一个主群列,由对应定理可知每个 C_i/G_{i+1} 是 G/G_{i+1} 的极小正规子群.

定理 5.2.17　有限群的极小正规子群是同构单群的直积.

推论 5.2.18　有限群的主因子是同构单群的直积.

5.3　可解群、幂零群与超可解群

上节我们引进了群的各种子群列的概念. 本节将利用子群列来研究群. 特别地,利用子群列来定义一类重要的群——可解群. Galois 引进群的概念,研究 5 次及 5 次以上方程的根式解问题,他证明了 $n(n\geqslant 5)$ 次方程有根式解当且仅当这个方程的 Galois 群是可解群. 本节将考察这类群及其相关群类.

5.3.1　可解群

定义 5.3.1　设 G 是群,令 $G^{(0)} = G$,对 $k \in \mathbf{N}$,定义 $G^{(k)}$ 为 $G^{(k-1)}$ 的**导群**,所得到的群列称为 G 的**导群列**.

对 $\forall k$,$G^{(k+1)}$ 为 $G^{(k)}$ 的特征子群;导群列是 G 的一个正规群列,而且导群列中任意相邻两群的商群都是交换群.

定义 5.3.2　称一个群为**可解群**,若它的导群列终止于 1.

由于一个群为交换群当且仅当其导群为 1,易见,交换群为可解群. 但是并不是所有群都可解,例如,A_5 不是可解群. 由于单群几乎没有正规子群,而可解群有许多正规子群,故可解群可认为是与单群相反的一个概念. 如果这一思想正确的话,那么就应该有很少的群既是单群又是可解群,而事实正是如此.

定理 5.3.1　可解单群为素数阶循环群.

证明　设 G 为可解单群,则 $G' \neq G$;又因为 G 是单群,故 $G' = 1$,从而 G 为交换群. 但由于交换单群的每个非单位元必是生成元,故 G 必为素数阶循环群.

下面给出可解群的一些性质:

定理 5.3.2　设 G 是群,则下列条件等价:

1) G 是可解群;

2) G 有一个正规群列,其相邻两群构成的商群为交换群;

3) G 有一个次正规群列，其相邻两群构成的商群为交换群.

证明　1)⇒2)⇒3)显然，故只需证 3)⇒1).

设 $G = G_0 \geqslant G_1 \geqslant \cdots \geqslant G_r = 1$ 是 G 的次正规群列，且任意相邻两群构成的商群为交换群，因为 $G_r = 1$，要证明 G 可解，只需证对 $\forall i, G^{(i)} \leqslant G$ 成立即可.

对 i 作归纳法. 由于 G/G_i 交换，则 $G^{(1)} \leqslant G_1$. 设 $i > 1$，由归纳假设得 $G^{(i-1)} \leqslant G_{i-1}$，则 $G^{(i)} = (G^{(i-1)})' \leqslant (G_{i-1})'$，由于 G_{i-1}/G_i 交换，则 $(G_{i-1})' \leqslant G_i$，故 $G^{(i)} \leqslant G_i$.

下面我们来考察一下与可解群相关的一些群的可解性.

定理 5.3.3　设 G 是群，则：

(1) 若 G 可解，$H \leqslant G$，则 H 可解；

(2) 若 G 可解，$N \lhd G$，则 G/N 可解；

(3) 若 $N \lhd G$，且 $N, G/N$ 可解，则 G 可解；

(4) 若 G, H 可解，则 $G \times H$ 可解.

由定理易知，对称群 S_2, S_3, S_4 均为可解群，而 S_5 不是可解群.

定理 5.3.4　有合成群列的群是可解群当且仅当它的所有合成因子为素数阶.

易见，存在既非单又不可解的群，如 S_5，也存在非交换的可解群，如 S_3 ($S_3 > A_3 > 1$). 没有合成群列的群有可能是可解群. 例如，无限交换群是可解群，但它没有合成群列.

定理 5.3.5　有限 p 群是可解群.

定义 5.3.3　设 p 是素数，Z_p 是 p 阶循环群，$n \in \mathbf{N}^+$，则 n 个循环群 Z_p 的直积 $Z_p \times Z_p \times \cdots \times Z_p$ 称为 p^n 阶**初等交换 p 群**.

容易验证，有限交换群 G 是初等交换 p 群的充要条件是 $\exp(G) = p$.

定理 5.3.6　有主群列的群是可解群当且仅当它的每个主因子为初等交换 p 群.

证明　设群 G 有主群列，若 G 的所有主因子为初等交换 p 群，则可加细群 G 的主群列，得到群 G 的一个合成群列，它的合成因子均为素数阶，又 G 是可解群. 反之，若 G 可解，设 H/K 为 G 的一个主因子，其中 $K \lhd H \lhd G$. 又 H/K 为可解群，而 H/K 的每个合成因子为素数阶，因而 H/K 有限. 由推论知，H/K 为同构于某一单群 S 的若干个群的直积. 由于 H/K 的每个合成因子必同构于 S，且这些因子为素数阶，故 H/K 是初等交换 p 群.

定理 5.3.7　有合成群列的群是可解群当且仅当它有一正规群列其相邻两群构成的商群为 p 群.

5.3.2　超可解群

定义 5.3.4　群 G 有一个有限的正规群列 $G = G_0 \geqslant G_1 \geqslant \cdots \geqslant C_r = 1$，其中每个商群 G_i/G_{i+1} 都是循环群，则称 G 为**超可解群**.

因为有限超可解群为可解群，且有限 p 群为超可解群. 但是并非所有的有限可解群均为超可解群.

例 5.3.1　S_4 有主群列 $S_4 > A_4 > K > 1$，其中 K 为 Klein 四元群，且 $S_4/A_4 \cong Z_2$，$A_4/K \cong Z_3$，$K \cong Z_2 \times Z_2$，则 S_4 是可解群，但不是超可解群.

定义 5.3.5　$N \lhd G$，N 在 G 中的**补子群**是指群 G 的子群 H，满足 $G = NH$，且$N \cap H = 1$.

定义 5.3.6　有限群 G 的子群 H 称为 **Hall 子群**，若 $|H|$ 与$[G:H]$互素.

定理 5.3.8(Schur-Zassenhaus 定理)　有限群的任意正规 Hall 子群均有补子群.

定理 5.3.9　设 G 是有限可解群，其阶为 mn，其中 $(m,n) = 1$，则：

1) G 有 m 阶子群；

2) G 的任意两个 m 阶子群均共轭；

3) 群 G 的阶整除 m 的子群包含在 G 的某一个 m 阶子群中.

这个定理对任意有限群不一定成立. 例如，$60 = 20 \times 3$ 阶群 A_5 没有 20 阶子群.

下面我们不加证明地给出有限群为可解群的一些著名定理.

定理 5.3.10(Burnside 定理)　设 p，q 是素数，a，b 为非负整数，则 $p^a q^b$ 阶群可解.

Burnside 定理是一个非常典型的结果，它告诉我们阶有 3 个因子的有限群不一定可解，且 A_5 就是一个反例.

定理 5.3.11(Feit-Thompson 定理)　奇阶群是可解群.

这个定理也称为“奇阶定理”，首先是由 Burnside 在 1911 年提出的猜想. 作为推论，易见奇阶单群只能是素数阶循环群.

定理 5.3.12　设 G 是有限群，若 G 有 m 阶子群，且$|H| = mn$，其中$(m,n) = 1$，则 G 是可解群.

定理 5.3.13　有限群 G 不可解当且仅当存在G 的阶互素的非单位元 x,y,z，使得 $xy = z$.

例 5.3.2　在 A_5 中取 $x = (1\ 2\ 3\ 4\ 5)$，$y = (1\ 2)(3\ 4)$，$z = (2\ 4\ 5)$，其中 $|x| = 5$，$|y| = 2$，$|z| = 3$，则 A_5 满足定理的条件，从而 A_5 不可解.

5.3.3　幂零群

下面介绍群论中的另一个重要群类，它介于交换群与可解群之间.

定义 5.3.7　群 G 的一个正规群列 $G = G_0 \geqslant G_1 \geqslant \cdots \geqslant G_r = 1$ 称为 G 的**中心群列**，若对每个 i，G_i/G_{i+1} 包含在 G/G_{i+1} 的中心里.

定义 5.3.8　若群 G 有一个中心群列，则称 G 为**幂零群**.

交换群 G 有中心群列 $G>1$，从而交换群是幂零群.

引理 5.3.14　有限 p 群是幂零群.

证明　设 P 为有限 p 群,则对 $|P|$ 作归纳法. 若 $|P|=p$,则 P 为交换群,从而幂零;若 $|P|>p$,令 $Z=Z(P)$,则 $Z\neq 1$. 由归纳假设知,P/Z 有中心群列

$$P/Z=P_0/Z\geqslant P_1/Z\geqslant\cdots\geqslant P_r/Z=1,$$

易见群列 $P=P_0\geqslant P_1\geqslant\cdots\geqslant P_r=Z\geqslant 1$ 是 P 的中心群列.

定理 5.3.15　设 G 是有限群,则下列陈述等价:

1) G 是幂零群;

2) 对 G 的任意子群 H,有 $N_G(H)>H$;

3) G 的每个 Sylow 子群在 G 中正规;

4) G 是其 Sylow 子群的直积;

5) G 的每个极大子群在 G 正规.

引理 5.3.16　循环 p 群 P 的每个非生成元必为 P 中某元的 p 次幂.

证明　P 恰有一个由 x^p 生成的指数为 p 的子群 Q,其中 x 为 P 的生成元. 若 $y\in Q$,则对某个整数 n,$y=(x^p)^n=(x^n)^p$,因而 Q 中每个元是 P 中某元的 p 次幂. 又 Q 是 P 的唯一的极大子群,P 的生成元素不能属于 P 的任何一个极大子群,故 Q 恰是由 P 的非生成元组成的集合,所以 P 的每个非生成元必为 P 中某元的 p 次幂.

定理 5.3.17　有限交换 p 群是循环 p 群的直积.

证明　设 P 为有限交换 p 群,对 $|P|$ 用归纳法. 设 $|P|>p$ 且结论对于阶小于 $|P|$ 的交换 p 群成立. 令 Q 是 P 的极大子群,则 $|P/Q|=p$. 由归纳假设,$Q=Q_1\times\cdots\times Q_s$,其中对 $\forall i$,Q_i 是 p^{a_i} 阶循环群. 设 $a_1\geqslant\cdots\geqslant a_s\geqslant 1$. 令 $x\in P$ 但 $x\notin Q$. 因为 $|P/Q|=p$,我们有 $x^p\in Q$,所以 $x^p=y_1\cdots y_s$,其中对 $\forall i$ 来说,$y_i\in Q_i$,若对某个 i 以及某个 $x_i\in Q_i$ 有 $y_i=x_i^p$,则 $(xx_i^{-1})^p=x^px_i^{-p}=x^py_i^{-1}=y_1\cdots y_{i-1}.1.y_{i+1}\cdots y_s$,但 $xx_i^{-1}\notin Q$,则存在 $x\in P-Q$ 使得 $x^p=y_1\cdots y_s$,其中 y_i 或为 Q_i 的生成元或为单位元.

若 $x^p\neq 1$,则 P 是 $<x>$ 与 Q 的直积,因而可设 $x^p\neq 1$,在此情况下,存在某个 i 使得 $y_i\neq 1$,令 $j(1\leqslant j\leqslant s)$ 是使得 $y_i\neq 1$ 的最小整数. 现在有 $x^p=y_j\cdots y_s$. 因为 P 是交换的,所以 x^p 的阶是 $y_j,\cdots,y_s$ 阶的最小公倍数 p^{a_j},因而 $|<x>|=p^{a_j+1}$. 令 $\overline{Q}\leqslant P$ 是除 Q_j 以外的所有 Q_i 的直积,则 $|\overline{Q}|=|Q|/|Q_j|=|Q|/p^{a_j}$.

若能证明 $<x>\cap\overline{Q}=1$,则 $<x>\overline{Q}$ 就是阶为 $p|Q|=|P|$ 的循环群的直积,因而 P 是循环群的直积. 设 $x^t\in<x>$,因 $x^p\in Q$,而对每个 $1\leqslant n<p$,$x^n\in Q$,故仅当 $p\mid t$ 时才有 $x^t\in Q$. 令 $t=mp$,其中 $0\leqslant m<p^{a_j}$,于是 $x^t=(x^p)^m=y_j^m\cdots y_s^m$. 因为 $m<p^{a_j}=|<y_j>|$,故 $y_j^m\neq 1$. 这说明 x^t 在 $Q=Q_1\times\cdots\times Q_s$ 的分解式中的第 j 个分量不为 1. 又由 $\overline{Q}$ 的作法知,$<x>\cap\overline{Q}=1$. 结论得证.

定理 5.3.18(有限交换群基本定理)　有限交换群是循环 p 群的直积.

定理 5.3.19　设 G 为 p^3 阶非交换群，则 G 同构于下列四种群之一：

1）当 $p \neq 2$ 时：

(i) $G = < x, y \mid y^{p^2} = x^p = 1, x^{-1}yx = y^{1+p} > \cong Z_{p^2} \times Z_p$；

(ii) $G = < x, y \mid x^p = y^p = z^p = 1, [x, y] = z, [x, z] = [y, z] = 1 > \cong (Z_p \times Z_p) \times Z_p$；

2）当 $p = 2$ 时：

(iii) $G = < x, y \mid y^4 = x^2 = 1, x^{-1}yx = y^2 > \cong Z_4 \times Z_2$（二面体群）；

(iv) $G = < x, y \mid x^4 = 1, x^2 = y^2, x^{-1}yx = y^2 >$（四元数群）.

5.3.4　幂零群与超可解群的关系

定理 5.3.20　幂零群是可解群.

例 5.3.3　S_3 没有中心群列. 否则，这样的群列中的倒数第二项必为 $Z(S_3)$ 的非平凡子群，显然这不可能，故 S_3 不是幂零群.

定理 5.3.21　有限幂零群是超可解群.

证明　设 G 为有限幂零群，$G = G_0 \geqslant G_1 \geqslant \cdots \geqslant C_r = 1$ 是 G 的一个中心群列. 若把这个中心群列加细为主群列，则它仍为 G 的一个中心群列. 由于包含于群 G 的中心里的子群必为 G 的正规子群，故 G 的主群列就是它的合成群列，因而列中每个商群为素数阶循环群，从而 G 为超可解群.

定理 5.3.22　超可解群的子群和商群是超可解的.

推论 5.3.23　超可解群满足极大条件.

超可解群是有限生成的，根据上述定理，它的子群也是有限生成的，所以满足极大条件.

定理 5.3.24　超可解群 G 具有正规群列 $G = H_0 \geqslant H_1 \geqslant \cdots \geqslant H_k = 1$，其中每个商群 H_i / H_{i+1} 是无限循环群或是素数阶循环群.

证明　设 $G = G_0 \geqslant G_1 \geqslant \cdots \geqslant G_r = 1$ 是正规群列，其中 G_{i-1}/G_i 是循环群，如果 G_{i-1}/G_i 有有限的阶 $p_1 p_2 \cdots p_s$，其中 $p_1, p_2 \cdots p_s$ 是素数（不必须是不同的），则 G_{i-1}/G_i 具有阶为 $p_1, p_1 p_2, \cdots, p_1 \cdots p_{s-1}$ 的唯一循环子群，而且这些都是特征子群，因此在 G_{i-1} 和 G_i 之间的 $s-1$ 个对应的子群在 G 内是正规的，而且相邻的群的商群是素数阶循环群. 用这个方法加细每个有限阶的商群 G_{i-1}/G_i，就得出定理中的正规群列，其中每个商群是无限循环群或素数阶循环群.

这个定理还可以进一步加强而按素数的大小来重新排列素数阶商群.

定理 5.3.25　超可解群的导出群是幂零的.

证明　设 $G = G_0 \geqslant G_1 \geqslant \cdots \geqslant G_r = 1$ 是 G 的正规群列，其中 G_{i-1}/G_i 是循环群，

记 $H_i = G' \cap G_i$,那么

$$G' = H_0 \geqslant H_1 \geqslant \cdots \geqslant H_r = 1$$

是正规序列,而且这序列中不同的项 k_i 组成群列

$$G' = K_0 \geqslant K_1 \geqslant \cdots \geqslant K_s = 1,$$

其中 K_{i-1}/K_i 是循环群.我们来判断这些由 K 组成 G' 的中心群列.

每个 K_i 是 G 中正规子群的交,因此在 G 内是正规的,这时在 G/K_i 内 K_{i-1}/K_i 是循环的正规子群,用 G/K_i 的元素作变形导出循环群 K_{i-1}/K_i 的自同构.然而循环群的自同构组成 Abel 群,所以 G/K_i 的两个元素导出 K_{i-1}/K_i 的可交换的自同构,于是任何两个元素的换位子 $x^{-1}y^{-1}xy$ 导出 K_{i-1}/K_i 的恒等自同构,这说明 K_{i-1}/K_i 属于 G'/K_i 的中心,因而这些 K 组成 G' 的中心群列,所以 G' 是幂零的.

我们已经介绍了几类重要的群:

$$\{循环群\} \subset \{交换群\} \subset \{幂零群\} \subset \{超可解群\} \subset \{可解群\}.$$

5.4　有限生成 Abel 群的结构

5.4.1　自由群

群论研究可以分成两个侧面,一方面是对给定的有背景的重要群,如各种对象的对称群、置换群、矩阵群(与几何、物理的联系)、有限单群、可解群(与解代数方程有联系)等等,讨论这些群的结构,以及它与其他群的关系(群的表示问题),另一方面是尽可能多地构造出一些新的群来,或者是借助于已知群去构造新群,或者根据需要去构造新群,通常称前者为群的结构理论和表示理论,而称后者为群的构造理论,当然这两者是互相联系的.

前面讨论过的子群、商群,都是从一个已知群获得新群的方法.

从一般线性群 $GL_n(F)$ 得到其子群 $SL_n(F)$,以及从 $SL_n(F)$ 得到 $PSL_n(F)$ 等等.

由两个已知群 K 和 H,由它们可以构造一个新的群,即它们的外直积 $H \times K$.由已知的两个群 H 和 K,还可用别的方法构造新群吗?

这样的群我们已见过,加群 $\mathbf{Z}$ 和循环群之间就有这种关系:任一循环群都是群 $\mathbf{Z}$ 的同态象.下面首先来推广这一结果.

定理 5.4.1　任意由 n 个元素生成的交换群 G 都是 $\mathbf{Z}^n$ 的同态象.

定义 5.4.1　称群 $\mathbf{Z}^n$ 为 n 阶**自由交换群**.

现在我们进一步研究某些群的结构,首先介绍约束条件最少的群——自由群以及用定义关系刻画群结构的方法.

定义 5.4.2　设 S 为任意集合，S 中有限个元素 $x_1, x_2, \cdots, x_n$ 连在一起叫做是一个**字**.

字 $x_1x_2\cdots x_n$ 和 $y_1y_2\cdots y_m$ 相等，如果 $n = m$ 且 $x_i = y_i, 1 \leqslant i \leqslant n$.

以 $\sum^*(S)$ 表示所有这样的字(包括空字 1) 组成的集合，在 $\sum^*(S)$ 中定义两个字的运算为：$(x_1x_2\cdots x_n)(y_1y_2\cdots y_m) = x_1\cdots x_ny_1\cdots y_m$.

对每个字 $\alpha \in \sum^*(S)$ 规定 $1\alpha = \alpha 1 = \alpha$，则这个运算显然满足结合律，从而 $\sum^*(S)$ 对上述运算形成一个幺半群，称为集合 S 上的**自由幺半群**.

集合 S 叫做 $\sum^*(S)$ 的**基**."自由"一词意味着 $\sum^*(S)$ 中除了含幺半群定义中的要求之外，没有任何其他拘束条件.

定义 5.4.3　如果 S 是有限集，则 $F(S)$ 叫做**有限生成自由群**.

特别地，当 $S = \{a\}$ 时，$F(S) = \langle a \rangle = \{a^n \mid n \in \mathbf{Z}\}$ 就是无限循环群，而当 $|S| \geqslant 2$ 时，$F(S)$ 是无限非 Abel 群.

定理 5.4.2　每个群都是自由群的商群. 每个有限生成群都是有限生成自由群的商群.

证明　设 G 为群. 取 G 的一个生成元系 $\sum$(例如可取 $\sum = G$)，定义集合 $S = \{X_a \mid a \in \sum\}$，并考虑映射 $f: F(S) \to G$，其中 $f(X_a) = a, f(X_a^{-1}) = a^{-1}$，对于 $A_i \in \sum \cup \sum^{-1} (1 \leqslant i \leqslant n)$，定义：$f(A_1\cdots A_n) = f(A_1)\cdots f(A_n)$. 这个映射是可以定义的，即不依赖于 $F(S)$ 中元素的不同表达方式，因为不同表达方式是由于插入或消去 $X_aX_a^{-1}$ 或 $X_a^{-1}X_a$ 造成的，而 $f(X_aX_a^{-1}) = aa^{-1} = 1, f(X_a^{-1}X_a) = a^{-1}a = 1$.

进一步，易知 f 是群同态，并且是满的，因为对每个生成元 $a \in \sum$，$a = f(X_a) \in \mathrm{Im}f$，从而 $G = \langle \sum \rangle = \mathrm{Im}f$. 根据同态基本定理，$G \cong F(S)/\mathrm{Ker}f$，即 G 同构于自由群 $F(S)$ 的商群. 如果 G 是有限生成的，令有限集 $\sum$ 是 G 的一个生成元系，则 $S = \{X_a \mid a \in \sum\}$ 也是有限集，从而 $F(S)$ 为有限生成自由群.

设 G 同构于自由群 $F(S)$ 的商群，$f: F(S)/K \cong G$，K 是 $F(S)$ 的正规子群，则 G 是由 $f(S) = \sum$ 生成的，进一步，对 K 中每个元素 α，G 中就有一个等式 $f(\alpha) = 1_G$，K 中有多少元素，G 中就相应有多少个关系. 如果 P 是 K 的一个子集，且 K 是 $F(S)$ 中包含 P 的最小正规子群(叫做由 P 生成的正规子群)，则 K 中每个元素均可由 P 在 $F(S)$ 中的全部共轭集合的元素运算出来，反映在群 G 中，G 的所有关系均可由 P 中元素给出的关系推导出来.

定义 5.4.4　我们把由 P 中元素给出的那些关系全体叫做群 G 的**定义关系集**，并且群 $G = \langle \sum \mid f(\alpha) = 1, \forall \alpha \in P \rangle$，这种刻画群的方式叫做群 G 的一个**表现**.

例如令 $S=\{a,b\}$，K 是 $F(S)$ 中的元素 a^3 和$(ab)^2$ 生成的正规子群，如果 $G\cong F(S)/K$，则 G 的结构可以写成 $G=<A,B\mid A^3=(AB)^2=1>$.

例 5.4.1 以 φ 表示关系集合为空集. $G=<S\mid\varphi>$ 即是以 S 为基的自由群，因为此时 $K=\{1\}$，$G\cong F(S)/\{1\}=F(S)$.

例 5.4.2 $Z_n\cong<a>/<a^n>=F(S)/<a^n>$，其中 $S=\{a\}$，因而 n 阶循环群的表现为 $Z_n=<a\mid,a^n=1>$.

例 5.4.3 正 $n(n\geqslant 3)$ 边形对称群 D_n 是 $2n$ 阶群，它有生成元 $<\sigma,\tau>$，其中 $\sigma^n=\tau^n=1$，$(\tau\sigma)^2=1$. 令 F 是以$\{a,b\}$ 为基的自由群，则有群的满同态 $f\colon F\to D_n$，$f(a)=\sigma$，$f(b)=\tau$，由同态基本定理可知 $D_n\cong F/\mathrm{Ker}f$.

例 5.4.4 令 $Q_8=<a,\ b\mid a^4=1,b^2=a^2,ba=a^3b>$. Q_8 中每个元素均可写成 $a^ib^j(0\leqslant i\leqslant 3,0\leqslant j\leqslant 1)$，从而 $|Q_8|\leqslant 8$. 但是我们有一个具体矩阵群 $G=<A,B>$，其中

$$A=\begin{pmatrix}0&1\\-1&0\end{pmatrix},B=\begin{pmatrix}0&\mathrm{i}\\\mathrm{i}&0\end{pmatrix}\quad(\mathrm{i}=\sqrt{-1}),$$

满足 $A^4=1$，$B^2=A^2$，$BA=A^3B$，可直接验证 $A^iB^j(0\leqslant i\leqslant 3,0\leqslant j\leqslant 1)$ 为 8 个不同的矩阵，因此 $|G|=8$. 然后可按例 5.4.3 的方法得出 $G\cong Q_8$，即 Q_8 为 8 阶非 Abel 群.

5.4.2 子群直积

定义 5.4.5 设 S 为任意集合，表现为 $F=<S\mid ba=ab,\forall a,b\in S>$ 的群叫做以 S 为基(或在 S 上) 的**自由 Abel 群**(除了交换性条件之外不再有任何关系).

由元素可交换知，F 中元素均可写成 $g=a_1^{n_1}a_2^{n_2}\cdots a_r^{n_r}(r\geqslant 0,\ n_i\in Z,\ n_i\neq 0,a_i\in S,1\leqslant i\leqslant r)$，其中 $a_1,a_2,\cdots,a_r$ 是 S 中不同元素，并且若不考虑前后次序，g 的这个表达方式是唯一的. 并且，每个(有限生成) 阿贝耳群均是(有限生成) 自由 Abel 群的商群.

为了进一步看清有限生成自由 Abel 群的结构，我们现在引进群的直积.

定义 5.4.6 设 $G_1,\cdots,G_n$ 是群，在集合的**直积**

$$G=G_1\times\cdots\times G_n=\{(g_1,\cdots,g_n)\mid g_i\in G_i,1\leqslant i\leqslant n\}$$

中定义运算

$$(g_1,\cdots,g_n)(g_1',\cdots,g_n')=(g_1g_1',\cdots,g_ng_n'),$$

设 G 和 K 是群，则 $G\times 1=\{(g,1)\mid g\in G\}$ 和 $1\times K=\{(1,k)\mid k\in K\}$ 是 $G\times K$ 的两个子群，并且 $G\times 1\cong G$，$1\times K\cong K$. $G\times 1$ 中元素和 $1\times K$ 中元素可交换，

$$G\times K=(G\times 1)(1\times K),(G\times 1)\cap(1\times K)=\{1\}.$$

引理 5.4.3　$H,K\leqslant G, H\cap K=\{1\}, G=HK$，对 $h\in H, k\in K, hk=kh$，则 $G\cong H\times K$.

证明　由 $G=HK$ 和 H 中元素与 K 中元素的交换性，可知 G 中每个元素均可表成 $g=hk, h\in H, k\in K$. 再由 $H\cap K=\{1\}$，g 的这个表达式是唯一的，于是我们可以定义 $f: G\to H\times K, hk\mapsto(h,k)$. 由上述可知这是一一对应的，并且 $f((hk)(h'k'))=f(hh',kk')=(hh',kk')=f(hk,h'k')$，从而 f 为同构，即 $G\cong H\times K$.

下面定理是判别一个群为某些子群直积的方法.

以后将 $G=G_1\times\cdots\times G_n$ 中元素 $(g_1,1,\cdots,1)$ 等同于 G_1 中元素 g_1，由此将 G_1 看成是 G 的正规子群. 类似地，$G_2,\cdots,G_n$ 也自然地看成 G 的正规子群，从而 G 的每个元素唯一地表成 $g=g_1\cdots g_n(g_i\in G_i)$.

定理 5.4.4　设 $G_1,\cdots,G_n$ 是 G 的正规子群，$n\geqslant 2$，则以下三条件是等价的：

1) $G=G_1\times\cdots\times G_n$；

2) G 中每个元素可以唯一表示成 $g=g_1\cdots g_n$，其中 $g_i\in G_i$；

3) $G=G_1\cdots G_n$，且对每个 $m, 1<m\leqslant n, (G_1G_2\cdots G_{m-1})\cap G_m=\{1\}$.

定义 5.4.7　如果 $G\cong Z^r$，则 r 叫做有限生成自由 Abel 群 G 的**秩**，记为 $\mathrm{rank}(G)$.

综合上述，我们证明了下面的结构定理：

定理 5.4.5　有限生成自由 Abel 群 G 同构于有限个无限循环群的直积，$G\cong\mathbf{Z}^r$，$r=\mathrm{rank}(G)\geqslant 1$. 两个这样的群 G 和 G' 同构 $\Leftrightarrow\mathrm{rank}(G)=\mathrm{rank}(G')$.

推论 5.4.6　设 S 和 S' 是有限生成自由 Abel 群 G 的两组基，则 $|S|=|S'|$.

5.4.3　有限生成 Abel 群的结构

我们把 Abel 群 A 中运算写成加法形式，幺元素为 0，元素 a 的逆是 $-a$，n 个 a 运算为 $a+a+\cdots+a=na$，有限阶元素 a 的阶为满足 $na=0$ 的最小正整数 n. $nA=\{na\mid a\in A\}$，直积则改叫做直和，并且写成 $A\oplus A'$，n 个 A 的直和仍表成 A^n.

定理 5.4.7　有限生成自由 Abel 群 F 的每个子群 $G(G\neq\{0\})$ 仍是有限生成自由 Abel 群，且 $\mathrm{rank}(G)\leqslant\mathrm{rank}(F)$.

定理 5.4.8　有限生成 Abel 群 A 均同构于 $\mathbf{Z}^r\oplus Z_{m_1}\oplus\cdots\oplus Z_{m_t}$，其中 $r,t\geqslant 0, 1<m_1\leqslant\cdots\leqslant m_t$ 且 $m_1\mid m_2\mid\cdots\mid m_t$.

设 A 是有限生成 Abel 群，以 A_t 表示 A 中有限阶元素全体. 如果 a 和 b 分别是 A 中阶数为 r 和 s 的元素，则 a^{-1} 和 ab 的阶分别是 r 和 $[r,s]$(r 和 s 的最小公倍数)，从而 A_t 是 A 的子群，易知 A_t 是有限 Abel 群.

定理 5.4.9　设 A 和 B 是有限生成 Abel 群，

1) 存在 A 的有限生成自由 Abel 子群 A_f，使得 $A=A_f\oplus A_t$；

2) 如果 $A=A_f\oplus A_t, B=B_f\oplus B_t$，其中 A_f 和 B_f 分别为 A 和 B 的有限生成自由

Abel 子群,则

$$A \cong B \Leftrightarrow \mathrm{rank} A_f = \mathrm{rank} B_f \text{ 且 } A_t \cong B_t.$$

定理 5.4.10　设 A 为有限 Abel 群, $A \neq \{0\}$,则

1) 存在 $1 < m_1 \mid m_2 \mid \cdots \mid m_t$, $(t \geqslant 1)$, 使 $A \cong Z_{m1} \oplus \cdots \oplus Z_{mt}$, $(m_1, \cdots, m_t)$ 由 A 唯一确定;

2) 存在一组正整数 $\{p_1^{s_1}, p_2^{s_2}, \cdots, p_k^{s_k}\}$,其中 $p_1, \cdots, p_k$ 为(不必不同的)素数, $s_1, \cdots, s_k$ 为正整数,使得 $A \cong Z_{p_1^{s_1}} \oplus \cdots \oplus Z_{p_k^{s_k}}$,且集合 $\{p_1^{s_1}, p_2^{s_2}, \cdots, p_k^{s_k}\}$ 由群 A 唯一决定.

定理中的 $\{m_1, \cdots, m_t\}$ 叫做 A 的不变因子, $\{p_1^{s_1}, \cdots, p_k^{s_k}\}$ 叫做 A 的初等因子. 对于有限生成 Abel 群 A, A 的不变因子和初等因子也分别叫做 A 的不变因子和初等因子. 综合上述,我们完成了有限生成 Abel 群的结构定理.

定理 5.4.11　两个有限生成 Abel 群同构⇔它们有相同的秩和初等因子⇔它们有相同的秩和不变因子.

特别地,两个有限 Abel 群同构⇔它们有相同的初等因子⇔它们有相同的不变因子.

第6章 环 与 域

本章研究代数结构的第二种类型——环.这些结构的原始模型是整数环 **Z**,这就是以半群的观点曾作为例子的(**Z**,+,0)和(**Z**,·,1).讨论这样的两个结构并把它们用分配律联系起来.环论与群论不同,群论在本质上仅有一个根源——就是研究关于积合成的双射的变换,而环论却是从许多专门理论当中抽象出来的,因为这个缘故,与群论比较,环论会显得不那么整齐和统一.但是以类似于产生变换群的方式同样产生环,就是 Abel 群的自同态环.

我们从环的各种类型的定义和例子来开始我们的讨论,环的这些类型是:整环、除环、交换环和域.随后我们研究理想、商环和同态这些基本概念,它们分别类似于正规子群、商群和群同态.在本章的后一部分,我们的注意力主要限制于交换环,首先考虑它们的某些扩张——交换整环的分式域、一个不定元的多项式环的构成和特性,随后我们考虑交换整环的因子分解的初等理论,特别是在数论方面的应用.

6.1 环论基础

6.1.1 环的定义和基本性质

定义 6.1.1 设 R 是一个非空集合,具有两种代数运算,加法("+")与乘法("·"),如果

1) $(R,+,0)$是一个 Abel 群;

2) $(R,\cdot,1)$是一个半群;

3) 分配律:$a(b+c)=ab+ac$,$(b+c)a=ba+ca$;

对 $\forall a,b,c\in R$ 都成立,则称 R 为一个**结合环**,记为$(R,+,\cdot)$.

结构$(R,+,0)$ 叫作 R 的**加法群**,而$(R,\cdot,1)$ 叫作 R 的**乘法半群**.

若环 R 的乘法运算适合交换律,则称 R 是**交换环**.

若在环 R 中,乘法有单位元,则称 R 是**有单位元环**.

设 G 是加群,规定 $ab=0$,$\forall a,b\in G$,则$(G,+,\cdot)$ 作成一个环,称为**零环**.

例 6.1.1 整数集 **Z** 关于数的加法、乘法作成一个环,称为**整数环**;偶数集 2**Z** 关于

数的加法、乘法也作成偶数环;同样 **Q**,**R**,**C** 关于数的加法、乘法都是环.

例 6.1.2 $\mathbf{Z}[i]=\{m+ni \mid m, n\in \mathbf{Z}\}$是一个数环,称为 Gaussian 整环.

数环 R 上全体 $m\times n$ 阶矩阵所组成的集合 $M_{m\times n}(R)$ 关于矩阵的加法、乘法作成一个环,称为 R 上的 $m\times n$ 阶全矩阵环.

一个加群 $E=\mathrm{Hom}(G, G)$ 是 G 的所有自同态所组成的集合,规定

$$\forall \sigma,\tau\in E, x\in G, (\sigma+\tau)(x)=\sigma(x)+\tau(x), (\sigma\cdot\tau)(x)=\sigma(\tau(x)),$$

则$(E, +, \cdot)$ 是一个环,称为 G 的**自同态环**.

$\mathbf{Z}_n=\{[0],[1],\cdots,[n-1]\}$ 关于加法运算:$[a]+[b]=[a+b]$,乘法运算:$[a][b]=[ab]$ 作成一个环(Z_n 称为模 n 的剩余类环).

环的一些初步性质如下:

4) $0+a=a+0=a, \forall a\in R$;

5) $a-a=a+(-a)=(-a)+a=0, \forall a\in R$;

6) $-(-a)=a, \forall a\in R$;

7) $a+b=c\Leftrightarrow b=c-a, \forall a,b,c\in R$;

8) $-(a+b)=-a-b, \forall a,b\in R$;

9) $m(na)=(mn)a; n(a+b)=na+nb, \forall n\in \mathbf{Z}, \forall a,b\in R$;

10) $(a-b)c=ac-bc, c(a-b)=ca-cb, \forall a,b,c\in R$;

11) $0a=a0=0, \forall a\in R$;

12) $(-a)b=a(-b)=-ab, (-a)(-b)=ab, \forall a,b\in R$;

13) $a(b_1+b_2+\cdots+b_n)=ab_1+ab_2+\cdots+ab_n, \forall a,b_i\in R$;

14) $(na)b=a(nb)=n(ab), \forall n\in \mathbf{Z}, \forall a,b\in R$.

定义 6.1.2 设 R 是一个环,$0\neq a\in R$,若存在 $0\neq b\in R$,使 $ab=0$,$(ba=0)$,则称 a 是 R 的一个左(右)**零因子**. 当 a 是左零因子,又是右零因子时,称 a 是 R 的零因子.

若环 R 有左零因子,则 R 必有右零因子. 因为若 a 是 R 的左零因子,则 $a\neq 0$,且存在 $0\neq b\in R$ 使得 $ab=0$,于是 b 是 R 的右零因子.

在整数环 **Z** 上的二阶全矩阵环 $M_2(\mathbf{Z})$ 中有零因子.

例 6.1.3 $[2],[3],[4],[6],[8],[9],[10]$是 $\mathbf{Z}_{12}$ 的所有零因子.

定义 6.1.3 环 R 中若 $ab=0$,则 $a=0$ 或 $b=0$($a,b\in R$),则称 R 是**无零因子环**.

整数环 **Z**、数域 F 上的一元多项式环 $F[x]$ 都是无零因子环,二阶全矩阵环 $M_2(\mathbf{Z})$、剩余类环 $\mathbf{Z}_6$ 都是有零因子环.

定理 6.1.1 设 R 是一个环，则 R 无左(右)零因子$\Leftrightarrow$在 R 中，乘法左、右消去律成立.

6.1.2 整环、除环和域

定义 6.1.4 一个有单位元、无零因子的交换环称为**整环**.

整数环 $\mathbf{Z}$、数域 F 上的一元多项式环 $F[x]$都是整环.

定义 6.1.5 设 R 是一个环，若

1) R 至少含有两个元素；

2) R 中每一个非零元都可逆，称 R 是一个**除环**(或体)，一个交换除环称为**域**.

除环具有下列性质：

1) 设环 R 至少含有两个元素，则 R 是除环 $\Leftrightarrow R$ 中全体非零元组成的集合 R^* 关于乘法作成一个群；

2) 除环没有零因子.

3) 一个至少含两个元素且没有零因子的有限环是除环，一个有限整环是域.

例 6.1.4 一个模 p 的剩余类环 $\mathbf{Z}_p$ 是域$\Leftrightarrow p$ 是素数.

证明 若 p 是素数，要证 $\mathbf{Z}_p$ 是域，因为 $\mathbf{Z}_p$ 是有单位元$[1]$的有限交换环，所以只要证 $\mathbf{Z}_p$ 无零因子即可. 任意$[a]$, $[b]\in\mathbf{Z}_p$，设$[a][b]=[0]$，则$[ab]=[0]$，于是，$p\mid ab$，因为 p 是素数，所以 $p\mid a$ 或 $p\mid b$，即$[a]=[0]$或$[b]=[0]$. 因此 $\mathbf{Z}_p$ 无零因子.

若 $p=1$，则 $|\mathbf{Z}_p|=1$，从而 $\mathbf{Z}_p$ 不是域. 若 p 是合数，则 $\mathbf{Z}_p$ 有零因子，从而 $\mathbf{Z}_p$ 不是域，因此，若 $\mathbf{Z}_p$ 是域，则 p 是素数.

下面给出一个非交换除环的例子.

例 6.1.5 $R=\{$所有复数对$(\alpha,\beta)\mid\alpha,\beta\in\mathbf{C}\}$. 这里$(\alpha_1,\beta_1)=(\alpha_2,\beta_2)$，当且仅当 $\alpha_1=\alpha_2$, $\beta_1=\beta_2$ 的时候. R 的加法和乘法是

$$(\alpha_1,\beta_1)+(\alpha_2,\beta_2)=(\alpha_1+\alpha_2,\beta_1+\beta_2);$$

$$(\alpha_1,\beta_1)(\alpha_2,\beta_2)=(\alpha_1\alpha_2-\beta_1\bar{\beta}_2,\alpha_1\beta_2+\beta_1\bar{\alpha}_2),$$

这里 $\bar{\alpha}$ 表示 α 的共轭数. 这样，R 是一个除环，称为四元数除环.

R 不是交换环，例如，$(0,i)=(i,0)(0,1)\neq(0,1)(i,0)=(0,-i)$.

6.1.3 环的特征

定义 6.1.6 设 R 是一个环，若存在最小正整数 n，对于所有 $a\in R$，都有 $na=0$，则称 n 是环 R 的**特征(数)**. 若这样的 n 不存在，则称环 R 的特征是零(或 ∞). 环 R 的特征记作 $\mathrm{ch}R$.

数域 F 的特征是 0,模 6 的剩余类环 Z_6 的特征是 6.

特征是一个很重要的概念,它对环的构造有着决定性的作用.

定理 6.1.2 · 在一个无零因子环 R 中,所有非零元的阶(加法)都相等.

证明　若 R 中每个非零元的阶都无限,则结论成立.

若存在 $0 \neq a \in R$,有 $|a| = n$,则 $na = 0$,存在 $0 \neq b \in R$,有 $(nb)a = b(na) = 0$,由于 R 是无零因子环,所以 $nb = 0$,从而 $|b| \leqslant n$,即 $|b| \leqslant |a$. 反之亦然,因此 $|b| = |a|$.

定理 6.1.3　设 R 是一个环,且 $\mathrm{ch}R = n > 0$,则

1) 当 R 是有单位元环时,n 是满足 $n \cdot 1 = 0$ 的最小正整数;

2) 当 R 是无零因子环时,n 是素数.

证明　1) 设 m 是使 $m1 = 0$ 的最小正整数,对 $\forall a \in R$,$ma = m(1a) = (m1)a = 0$,有 $n = m$.

2) 显然 $n \neq 1$. 又若 n 不是素数,则 $n = n_1 n_2$,$1 < n_1, n_2 < n$,对于 $0 \neq a \in R$,$|a| = n$,从而 $n_1 a \neq 0$,$n_2 a \neq 0$,但是 $(n_1 a)(n_2 a) = (n_1 n_2)a = na = 0$,这与 R 无零因子矛盾,则 n 是素数.

推论 6.1.4　域 F 的特征或是素数,或是零.

6.1.4　子环

定义 6.1.7　设 R 是一个环,$\varnothing \neq S \subseteq R$,若 S 关于 R 的加法、乘法作成环,则称 S 是 R 的一个**子环**,记作 $S \leqslant R$.

类似地,可以定义**子整环**、**子除环**、**子域**等概念.

例如,对于任意一个环 R,都有两个子环:$\{0\}$ 与 R. 这两个子环称为 R 的**平凡子环**. 若 $S \leqslant R$,且 $S \neq \{0\}$,$S \neq R$,则称 S 是 R 的**真子环**,记作 $S < R$.

由子环的定义与子群、子半群的判别方法,我们得到:

定理 6.1.5　1) R 是环,$\varnothing \neq S \subseteq R$,则 $S \leqslant R \Leftrightarrow \forall a, b \in S$,有 $a - b, ab \in S$;

2) S 是 R 的子除环(域) $\Leftrightarrow \forall a, b \in S$,有 $a - b, ab^{-1} \in S$.

例 6.1.6　在 $R[x]$ 中,R 是 $R[x]$ 的一个子环;所有常数项为零的一元多项式组成的集合 $S = \{a_1 x + a_2 x^2 + \cdots + a_n x^n \mid a_i \in R, n \in N\}$ 也是 $R[x]$ 的一个子环.

例 6.1.7　在数域 F 上的 2 阶全矩阵环 $M_2(F)$ 中,令

$$S_1 = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \middle| a, b, d \in F \right\}; S_2 = \left\{ \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \middle| a, b \in F \right\}; S_3 = \left\{ \begin{pmatrix} a & 0 \\ c & 0 \end{pmatrix} \middle| a, c \in F \right\};$$

$$S_4 = \left\{ \begin{pmatrix} a & 0 \\ 0 & d \end{pmatrix} \middle| a, d \in F \right\}; S_5 = \left\{ \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} \middle| a \in F \right\},$$

则 S_1, S_2, S_3, S_4, S_5，以及 $M_2(F)$ 本身都是 $M_2(F)$ 的子环.

例 6.1.8 设 R 是一个环，令 $C(R)=\{c\in R \mid cx=xc, \forall x\in R\}$，则 $C(R)$ 是 R 的交换子环，$C(R)$ 称为 R 的中心.

(1) 在交换性上：

1) 若只是交换环，则 S 也是交换环，

2) 当 S 是交换环时，R 未必是交换环. 例如，实数域 R 的 2 阶全矩阵环 $M_2(R)$ 是非交换环，而其中心 $C(M_2(R))$ 是交换环.

(2) 在有无零因子上：

1) 若 R 是无零因子环，则 S 也是无零因子环；

2) 当 S 是无零因子环时，R 未必是无零因子环. 例如，Z_{12} 是有零因子环，而其子环 S_3 是无零因子环.

(3) 在有无单位元上：

1) 若 R 有单位元，S 可以没有单位元. 例如，整数环 $\mathbf{Z}$ 有单位元 1，但其子环 $2\mathbf{Z}$(偶数环) 没有单位元；又如，全矩阵环 $M_2(F)$ 有单位元，但其子环 S_3, S_4 没有单位元.

2) 若 S 有单位元，只可以没有单位元. 例如，S_3 没有单位元，但其子环 S_5 有单位元.

3) 若 R 与 S 都有单位元，它们的单位元可以不同. 例如，在例 1.7 中，$M_2(F)$ 有单位元 $I_2=\begin{pmatrix}1&0\\0&1\end{pmatrix}$，但其子环 S_5 有单位元$\begin{pmatrix}1&0\\0&0\end{pmatrix}$.

定理 6.1.6 $\varnothing\neq T\subseteq R$，令 $S=\{\sum\pm x_1x_2\cdots x_n \mid x_i\in T, n\in\mathbf{N}\}$，则 $S\leqslant R$.

证明 由 $T\neq\varnothing$ 知 $S\neq\varnothing$. 又 $\forall a,b\in S$，设

$$a=\sum\pm x_1x_2\cdots x_n, b=\sum\pm y_1y_2\cdots y_m,$$

其中 $x_i, y_j\in T, n, m\in\mathbf{N}$，则

$$a-b=\sum\pm x_1x_2\cdots x_n-\sum\pm y_1y_2\cdots y_m\in S,$$

$$ab=(\sum\pm x_1x_2\cdots x_n)(\sum\pm y_1y_2\cdots y_m)=\sum\pm x_1x_2\cdots x_n y_1y_2\cdots y_m\in S,$$

因此，$S\leqslant R$.

定理所构造的子环 S 称为由 T 所生成的子环，记作$[T]$. 若 $T=\{t_1,t_2,\cdots,t_n\}$ 是有限集，则称$[T]$ 是有限生成的，并记作$[t_1,t_2,\cdots,t_n]$.

特别地，$[t]=\{\sum_{i=1}^{m}n_it^i \mid n_i\in\mathbf{Z}, m\in\mathbf{N}\}$.

定理 6.1.7 设 R 是环，$\varnothing\neq T\subseteq R, M=\{S_i \mid T\subseteq S_i\leqslant R, i\in I\}$ 是 R 的所有包含 T 的子环族，则$[T]=\bigcap_{i\in I}S_i$.

证明　因为 $T\subseteq[T]$,所以$[T]\in M$,从而 $\bigcap_{i\in I}S_i\subseteq[T]$;又因为 $T\subseteq S_i,i\in I$,所以 $T\subseteq\bigcap_{i\in I}S_i$,从而$[T]\subseteq\bigcap_{i\in I}S_i$,因此$[T]=\bigcap_{i\in I}S_i$.

$[T]$ 是 R 中包含 T 的最小子环.

6.2　理想与商环

6.2.1　环的同态

定义 6.2.1　设 R 与 R' 都是环,f 是 R 到 R' 的映射,若 f 保持运算,即 $\forall x,y$ 有 $f(x+y)=f(x)+f(y)$,$f(xy)=f(x)f(y)$,称 f 是 R 到 R' 的**同态映射**.

若同态 f 是满射,则称 f 是**满同态**,并称 R 与 R' **同态**,记作 $R\sim R'$;若同态 f 是单射,则称 f 是**单同态**;若同态 f 是双射,则称 f 是**同构**,并称 R 与 R' **同构**,记作 $R\cong R'$.

特别地,R 与 R 的同态又称为 R 的**自同态**,R 与 R 的同构又称为 R 的**自同构**.

设 R 与 R' 都是环,令 $f(x)=0'$,$\forall x\in R$,则 f 是 R 到 R' 的映射,称为**零同态**.

例 6.2.1　设 $R=\mathbf{Z}$,$R'=\mathbf{Z}_m$,令 $f(n)=[n]$,$\forall n\in\mathbf{Z}$,则 f 是 R 到 R' 的一个满同态.

例 6.2.2　设 $\mathbf{Z}[x]$是整数环 $\mathbf{Z}$ 上的一元多项式环,令 $\varphi(f(x))=f(0)$,$\forall f(x)\in\mathbf{Z}[x]$,则 φ 是 $\mathbf{Z}[x]$ 到 Z 的一个满映射,且 $\forall f(x),g(x)\in\mathbf{Z}[x]$,有

$$\varphi(f(x)+g(x))=f(0)+g(0)=\varphi(f(x))+\varphi(g(x)),$$
$$\varphi(f(x)g(x))=f(0)g(0)=\varphi f(x)\varphi g(x),$$

所以 φ 是 $\mathbf{Z}[x]$ 到 $\mathbf{Z}$ 的满同态,从而 $\mathbf{Z}[x]\sim\mathbf{Z}$.

设 f 是环 R 到环 R' 的同态,则有下列性质.

1) 若 0 是 R 的零元,则 $f(0)$ 是 R' 的零元,$\forall a\in R$,$f(-a)=-f(a)$;

2) 若 $S\leqslant R$,则 $f(S)\leqslant R'$;若 $S'\leqslant R'$,则 $f^{-1}(S')\leqslant R$.

当 $f:R\to R'$ 是满同态时,R 与 R' 在是否可交换、有无零因子、有无单位元等性质上有一定的联系,但并不完全一致,若 $R\cong R'$,则环 R 与环 R' 的代数性质完全一致.

(1) 在交换性上:

1) 若 R 是交换环,则 R' 也是交换环.

对 $\forall x',y'\in R'$,存在 $x,y\in R$,使 $f(x)=x'$,$f(y)=y'$,则 $x'y'=f(x)f(y)=f(xy)=f(yx)=f(y)f(x)=y'x'$.

2) 当 R' 是交换环时,R 未必是交换环.例如,在例 6.1.7 中的 S_1 是非交换环,而其在 f_1,f_2 下的同态像 S_4,S_5 都是交换环.

(3) 在有无单位元上:

1) 若 R 有单位元 1,则 R' 有单位元 $f(1)$.

对 $\forall x'\in R$,存在 $x\in R$,使 $f(x)=x'$,所以 $x'f(1)=f(x)f(1)=f(x1)=$

$f(x)=x'$，同理，$f(1)x'=x'$，因此 $f(1)$ 是 R' 的单位元.

2）当 R' 有单位元时，R 未必有单位元. 例如，在例 6.1.7 中的 S_2 没有单位元，而其在 f_4 下的同态像 S_5 有单位元.

(3) 在有无零因子上：

1）当 R 是无零因子环时，R' 未必是无零因子环. 例如，整数环 $\mathbf{Z}$ 是无零因子环，而其在 f 下的同态像 $\mathbf{Z}_m$，当 m 是合数时是有零因子环.

2）当 R' 是无零因子环时，R 未必是无零因子环. 例如，$R=\{(a,b)\mid a,b\in\mathbf{Z}\}$ 有零因子：$(1,0)$，$(0,1)$，而其在 f 下的同态像 $\mathbf{Z}$ 是无零因子环.

定理 6.2.1　设 $R\cong R'$，则 R 是整环(除环，域)$\Leftrightarrow R'$ 是整环(除环，域).

定理 6.2.2　设 $f:R\sim R'$，$g:R'\sim R$，则 f 与 g 的合成 $g\cdot f:R\sim R''$.

证明　显然 $g\cdot f$ 是 R 到 R'' 的映射，又对 $\forall x,y\in R$，有

$$(g\cdot f)(x+y)=g(f(x+y))=g(f(x)+f(y))$$
$$=g(f(x))+g(f(y))=(g\cdot f)(x)+(g\cdot f)(y),$$
$$(g\cdot f)(xy)=g(f(xy))=g(f(x)f(y))$$
$$=g(f(x))g(f(y))=(g\cdot f)(x)(g\cdot f)(y),$$

因此 $g\cdot f$ 是同态.

设 f 是环 R 到环 R' 的同态，则 f 可以看作加群 $(R,+)$ 到加群 $(R',+)$ 的同态，我们把 f 关于加群的同态核 $\mathrm{Ker}f$ 作为关于环的同态核.

定义 6.2.2　设 $f:R\sim R'$，称 $\mathrm{Ker}f=\{x\in R\mid f(x)=0'\}$ 是 f 的**同态核**.

定理 6.2.3　设 $f:R\sim R'$，0 是 R 的零元，则 f 是单同态 $\Leftrightarrow\mathrm{Ker}f=\{0\}$.

6.2.2　理想

设 $(R,+,\cdot)$ 是一个环，$(A,+)$ 是 $(R,+)$ 的一个子加群，则有商群 $R/A=\{x+A\mid x\in R\}$，其加法是 $(x+A)+(y+A)=(x+y)+A$. 现在要使 R/A 成为一个环，再定义一个乘法，最自然的办法是令 $(x+A)(y+A)=xy+A$.

然而 R/A 中的元素是陪集，这就要考虑具备什么条件的 A 才使这个规定与代表元的选取无关？即如果 $x_1+A=x+A$，$y_1+A=y+A$，要求 $x_1y_1+A=xy+A$，即 $x_1y_1-xy\in A$.

由上式得 $x_1=x+a$，$y_1=y+a'$，$a,a'\in A$，于是 $x_1y_1-xy=xa'+ay+aa'$，从而要求 $xa'+ay+aa'\in A$，为此只要 $xa',ay\in A$，$\forall x,y\in R$，$a,a'\in A$.

定义 6.2.3　设 $(R,+,\cdot)$ 是一个环，$(A,+)$ 是 $(R,+)$ 的一个子加群，

1）若 $\forall r\in R$，有 $ra\in A$，则称 A 是 R 的**左理想**；

2）若 $\forall r\in R$，$a\in A$，有 $ar\in A$，则称 A 是 R 的**右理想**；

3) 若A既是R的左理想，又是R的右理想，则称A是R的(双侧)**理想**，记作$A \lhd R$. 若$A \lhd R$，且$A \neq R$，则称A是R的**真理想**. 由定义知理想一定是子环.

任意一个环$R \neq 0$都有两个理想：$\{0\}$(称为**零理想**)与R(称为**单位理想**).

例 6.2.3 设R是交换环，$a \in R$，则$A = \{ar \mid r \in R\}$是R的一个理想.

设R是一个数环，则$\{a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x \mid a_i \in R, n \in \mathbf{N}\}$是$R[x]$的一个理想.

设R是一个环，$A_1 \lhd R$，$A_2 \lhd R$，则$A_1 \cap A_2 \lhd R$.

设R是一个环，$A_1 \lhd R$，$A_2 \lhd R$，则$A_1 + A_2 = \{a_1 + a_2 \mid a_1 \in A_1, a_2 \in A_2\} \lhd R$.

定义 6.2.4 只有零理想与单位理想的环称为**单环**.

例 6.2.4 除环R是单环.

证明 设$0 \neq A \lhd R$，则存在$0 \neq a \in A$，于是$1 = a^{-1}a \in A$，从而$\forall r \in R$，有$r = r1 \in A$. 因此$A = R$，即R是单环.

设R是一个环，$\varnothing \neq T \subseteq R$，$M = \{A_i \mid T \subseteq A_i \lhd R, i \in I\}$是$R$中所有包含$T$的理想族，则称$\bigcap_{i \in I} A_i$是由$T$所生成的理想，记作$(T)$，称$T$的元素是$(T)$的生成元.

显然(T)是R中包含T的最小理想. 由一个元a生成的理想(a)称为**主理想**.

定理 6.2.4 R是环，$a \in R$，则$(a) = \{\sum x_i a y_i + sa + at + na \mid x_i, y_i, s, t \in R, n \in \mathbf{N}\}$.

推论 6.2.5 设R是一个环，$a \in R$，

1) 若R是有单位元环，则$(a) = \{\sum x_i a y_i \mid x_i, y_i \in R\}$；

2) 若R是交换环，则$(a) = \{ra + na \mid r \in R, n \in \mathbf{Z}\}$；

3) 若R是有单位元的交换环，则$(a) = \{ra \mid r \in R\}$.

例 6.2.5 求证，整数环$\mathbf{Z}$的每一个理想都是主理想.

证明 因为Z是有单位元的交换环，所以$\forall n \in \mathbf{Z}$，$(n) = \{rn \mid r \in \mathbf{Z}\}$.

现设$A \lhd \mathbf{Z}$，若$A = \{0\}$，则$A = (0)$. 若$A \neq \{0\}$，则A中存在最小正整数n，于是$(n) \subseteq A$. 又$\forall a \in A$，设$a = nq + t$，$0 \leqslant t < n$，因为$a \in A$，$n \in (n) \subseteq A$，所以$t = a - nq \in A$. 由n的最小性，得$t = 0$，从而$a = nq \in (n)$，即$A \subseteq (n)$.

例 6.2.6 确定一元多项式环$\mathbf{Z}[x]$的理想$(2, x)$，证明$(2, x)$不是主理想.

证明 因为$\mathbf{Z}[x]$是有单位元的交换环，所以

$$\begin{aligned}(2, x) &= \{2f_1(x) + xf_2(x) \mid f_1(x), f_2(x) \in \mathbf{Z}[x]\} \\ &= \{2a_0 + xf(x) \mid a_0 \in \mathbf{Z}, f(x) \in \mathbf{Z}[x]\},\end{aligned}$$

即$(2, x)$是由$\mathbf{Z}[x]$中常数项为偶数的多项式所组成的.

若$(2, x)$是主理想，则存在$p(x) \in \mathbf{Z}[x]$，使$(2, x) = (p(x))$，于是$2 = p(x)q(x)$，$x = p(x)h(x)$，$q(x), h(x) \in \mathbf{Z}[x]$. 而由$2 = p(x)q(x)$，得$p(x) = c \in \mathbf{Z}$. 再由$x =$

$p(x)h(x)=ch(x)$ 得 $c=\pm 1$. 于是 $\pm 1=p(x)\in(2,x)$,但这与 ± 1 不属于 $(2,x)$ 矛盾. 因此,$(2,x)$ 不是主理想.

6.2.3 同态基本定理

定义 6.2.5 设 $A\lhd R$,在商群 $(R,+)/(A,+)=\{[x]\mid x\in R\}=\{x+A\mid x\in R\}$ 中再规定:$[x][y]=[xy]$,$\forall[x],[y]\in R/A$,则$(R/A,+,\cdot)$是一个环. R/A 称为 R 关于 A 的**商环**,或**剩余类环**,$[x]=x+A$ 称为 R 模 A 的**剩余类**.

若 R 是交换环,则 R/A 也是交换环;若 R 是有单位元1的环,则 R/A 有单位元$[1]$.

例 6.2.7 求一元多项式环 $R[x]$关于主理想(x)的商环 $R[x]/(x)$.

解 $R[x]/(x)=\{[f(x)]\mid f(x)\in R[x]\}$,且 $[f(x)]=[g(x)]\Leftrightarrow(x)-g(x)\in(x)\Leftrightarrow x\mid f(x)-g(x)$,所以 $\forall f(x)\in R[x]$,设 $f(x)=xq(x)+a,a\in R$,则$[f(x)]=[a]$,因此 $R[x]/(x)=\{[a]\mid a\in R\}$.

定理 6.2.6 一个环 R 与它的每一个商环 R/A 同态,称为环的**自然同态**,设 ψ 是环 R 到商环 R/A 的**自然同态**,则 $\mathrm{Ker}\psi=A$.

定理 6.2.7(同态基本定理) 设 f 是环 R 到环 R'的同态,则

1) $\mathrm{Ker}f\lhd R$; 2) $R/\mathrm{Ker}f\cong\mathrm{Im}f$.

证明 1) 已知 $\mathrm{Ker}f$ 是$(R,+)$的子加群,又 $\forall a\in\mathrm{Ker}f,r\in R$ 有 $f(ra)=f(r)f(a)=f(r)0'=0'$,$f(ar)=f(a)f(r)=0'f(r)=0'$,从而 $ra,ar\in\mathrm{Ker}\,f$,因此 $\mathrm{Ker}f\lhd R$.

2) 存在双射 φ: $a+\mathrm{Ker}f\rightarrow f(a)$,且保持加法运算. 又 $\forall a+\mathrm{Ker}f,b+\mathrm{Ker}f\in R/\mathrm{Ker}f$,有 $\varphi[(a+\mathrm{Ker}f)\cdot(b+\mathrm{Ker}f)]=\varphi(ab+\mathrm{Ker}f)=f(ab)=f(a)f(b)=\varphi(a+\mathrm{Ker}f)\cdot\varphi(b+\mathrm{Ker}f)$,从而 φ 保持乘法运算,因此 φ 是同构,从而 $R/\mathrm{Ker}\,f\cong\mathrm{Im}f$.

例 6.2.8 设 f 是环 R 到环 R'的满同态,则 $R/\mathrm{Ker}f\cong R'$.

例 6.2.9 求证,$R[x]/(x)\cong R$.

证明 令 $\varphi(f(x))=f(0)$,f: $R\sim R'$,$\forall f(x)\in R[x]$,易证 φ 是 $R[x]$ 到 R 的满同态,且 $\mathrm{Ker}\varphi=\{f(x)\in R[x]\mid\varphi(f(x))=0\}=\{f(x)\in R[x]\mid f(0)=0\}=\{xg(x)\mid g(x)\in R[x]\}=(x)$,于是,由同态基本定理得,$R[x]/(x)\cong R$.

定理 6.2.8(第一同构定理) 设 f 是环 R 到环 R' 的满同态,$A'\lhd R'$,$A=f^{-1}(A')$,则 $A\lhd R$,并且 $R/A\cong R'/A'$.

推论 6.2.9 设 f 是环 R 到环 R'的满同态,若 $A\lhd R$,则 $f(A)\lhd R'$.

6.2.4 素理想与极大理想

在整数环 $\mathbf{Z}$ 中,素数有两种刻画方法. 通常定义为:素数是一个大于1的整数,且除

1 和 p 自身外无其他正因数;另外也可以定义为:素数是一个大于1的整数,且 $\forall a,b \in \mathbf{Z}, p \mid ab \Rightarrow p \mid a$ 或 $p \mid b$. 如果用理想来描述,那么素数的第一个特征为:(p) 是 $\mathbf{Z}$ 的真理想,且若有 $\mathbf{Z}$ 的理想 A 满足 $(p) \subset A$,则 $A=\mathbf{Z}$;而素数的第二个特征为:若 $ab \in (p)$,则 $a \in (p)$ 或 $b \in (p)$. 把素数的这两个性质推广到一般交换环上,就得到极大理想与素理想两个概念.

定义 6.2.6 设 R 是交换环,P 是 R 的一个理想,若 $\forall a,b \in R, ab \in P \Rightarrow a \in P$ 或 $b \in P$,则称 P 是 R 的**素理想**.

显然,单位理想是素理想. 又当 R 是无零因子交换环时,零理想也是素理想. 但是,在 R 是有零因子时,由于存在 $a \neq 0, b \neq 0$,使 $ab=0$,即 $ab \in (0)$,且 a 不属于 (0),b 不属于 (0),从而零理想不是 R 的素理想.

例 6.2.10 设 p 是素数,则 (p) 是整数环 $\mathbf{Z}$ 的素理想.

定理 6.2.10 设 $P \lhd R$,则 P 是 R 的素理想 $\Leftrightarrow R/P$ 是整环.

证明 设 P 是 R 的素理想,若在 R/P 中有 $[a][b]=0$,即 $[ab]=[0]$,则 $ab \in P$. 于是 $a \in P$ 或 $b \in P$,即 $[a]=[0]$ 或 $[b]=[0]$. 因此 R/P 无零因子. 又因为 R 是有单位元的交换环,从而其商环 R/P 也是有单位元的交换环,因此 R/P 是整环.

设 R/P 是整环,$\forall a,b \in R$,若 $ab \in P$,则 $[ab]=[0]$,即 $[a][b]=[0]$. 因为 R/P 是无零因子环,所以 $[a]=[0]$ 或 $[b]=[0]$,即 $a \in P$ 或 $b \in P$,因此 P 是 R 的素理想.

例 6.2.11 (x) 是整数环 $\mathbf{Z}$ 上一元多项式环 $\mathbf{Z}[x]$ 的素理想.

证明 用上例同样方法,容易证明,$\mathbf{Z}[x]/(x) \cong \mathbf{Z}$. 而 $\mathbf{Z}$ 是整环,于是 $\mathbf{Z}[x]/(x)$ 也是整环. 因此,(x) 是 $\mathbf{Z}[x]$ 的素理想.

定义 6.2.7 设 M 是环 R 的一个真理想,若对于 R 的理想 $N, M \subset N$,有 $N=R$,则称 M 是 R 的**极大理想**.

(x) 是实数域 R 上一元多项式环 $R[x]$ 的极大理想 .

例 6.2.12 设 p 是素数,则 (p) 是整数环 $\mathbf{Z}$ 的极大理想.

证明 1 不属于 (p),从而 $(p) \neq \mathbf{Z}$. 设有 $\mathbf{Z}$ 的理想 N,使 $(p) \subset N$,则存在 $q \in N \backslash (p)$. 因为 q 不属于 (p),所以 p 不整除 q. 又由于 q 是素数,从而 $(p,q)=1$,即存在 $s, t \in \mathbf{Z}$,使 $sp+tq=1$. 因为 $p \in (p) \subset N, q \in N$,所以 $1 \in N$,从而 $N=\mathbf{Z}$. 因此 (p) 是整数环 $\mathbf{Z}$ 的极大理想.

一个环可以有多个极大理想,也可以没有极大理想,然而一个有单位元的环一定含有极大理想.

定理 6.2.11 设 M 是有单位元的交换环 R 的一个理想,则 M 是 R 的极大理想的充分必要条件是 R/M 是域.

证明 设 M 是 R 的极大理想. 由于 R 是有单位元的交换环,于是 R/M 也是有单位元的交换环,从而只要证 $0 \neq [a] \in R/M$ 在 R/M 中可逆. 令 $N=\{m+ax \mid m \in M,$

$x \in R\}$,容易证明 N 是 R 的一个理想,且 $M \subseteq N$,但是 a 不属于 M,$a \in N$,而 M 是 R 的极大理想,从而 $N = R$,于是 $1 \in N$. 存在 $x \in R, m \in M$,使 $1 = m + ax$. 这样 $[1] = [m + ax] = [ax] = [a][x]$,即 $[a]$ 是 R/M 的可逆元,因此 $R - M$ 是域.

设 R/M 是域,N 是 R 的理想,且 $M \subset N$,欲证 $N = R$.

由于 $M \subset N$,存在 $a \in N \backslash M$,即 $[0] \neq [a]$. 因 R/M 是域,所以存在 $x \in R$,使 $[a][x] = [1]$,即 $[ax] = [1]$,从而 $1 - ax \in M \subset N$. 由于 $a \in N$,N 是 R 的理想,于是 $ax \in N$,从而 $1 \in N$,所以 $N = R$. 因此 M 是 R 的极大理想.

推论 6.2.12 在有单位元的交换环 R 中,极大理想一定是素理想.

例 6.2.13 在一元多项式环 $\mathbf{Z}[x]$ 中,$(2, x)$ 是一个极大理想,而 (x) 不是极大理想.

6.2.5 商域

上一节我们介绍了利用极大理想构造域的方法,本节将介绍另一个通过环的扩充构造域的方法. 为此,先证明一个常用定理.

定理 6.2.13(挖补定理) 设 $S \leqslant R, S \cong S', S' \cap R \neq \varnothing$,则存在 S' 的扩环 R' 使 $R \cong R'$.

证明 设 S 到 S' 的同构为 φ. 令 $R' = S' \cup (R \backslash S)$,一个 R 到 R' 的映射

$$f(r) = \begin{cases} \varphi(r), r \in S, \\ r, r \in R \backslash S, \end{cases}$$

$\forall r' \in R'$,若 $r' \in S'$,由于 φ 是 S 到 S' 的同构映射,从而存在 $r \in S$,使 $\varphi(r) = r'$,即 $f(r) = r'$. 若 $r' \in R \backslash S$,则 $f(r') = r'$. 所以 f 是满射.

下面证 f 是单射. $\forall r_1, r_2 \in R$,若 $r_1 \neq r_2$,当 r_1, r_2 同属于 $R \backslash S$ 时,由于 $f(r_1) = r_1$, $f(r_2) = r_2$,所以 $f(r_1) \neq f(r_2)$. 当 r_1, r_2 同属于 S 时,因为 ψ 是同构映射,所以 $\varphi(r_1) \neq \varphi(r_2)$,于是 $f(r_1) \neq f(r_2)$. 当 r_1, r_2 一个属于 S,另一个属于 $R \backslash S$ 时,设 $r_1 \in S, r_2 \in R \backslash S$,则 $f(r_1) = \varphi(r_1) \in S, f(r_2) = \varphi(r_2) \in R$,由于 $S' \cap R = \varnothing$,所以 $f(r_1) \neq f(r_2)$.
因此 f 是 R 到 R' 的双射. 现通过 f,在 R' 中规定两种代数运算: $\forall r_1', r_2' \in R'$,存在 r_1, $r_2 \in R$,使 $f(r_1) = r_1', f(r_2) = r_2'$,令 $r_1' \oplus r_2' = f(r_1 + r_2), r_1' r_2' = f(r_1 r_2)$. 这样 f 是从 R 到 R' 的保持运算的双射,于是 R' 是一个环,从而 $R \cong R'$.

最后证明 S' 是 R' 的子环,这只要证明 S' 中的元素在 S' 中的运算与在 R' 中的运算一致. $\forall r_1', r_2' \in S', \exists r_1, r_2 \in S$,使 $\varphi(r_1) = r_1', \varphi(r_2) = r_2'$,于是

$$r_1' + r_2' = \psi(r_1) + \psi(r_2) = f(r_1) + f(r_2) = f(r_1 + r_2) = r_1' \oplus r_2',$$

$$r_1' r_2' = \psi(r_1)\psi(r_2) = f(r_1) f(r_2) = f(r_1 r_2) = r_1' r_2',$$

因此 R' 是一个符合条件的环.

我们知道整数环 **Z** 是一个整环，不是域. 但是将 **Z** 扩充得到了有理数域. 一般的环是否也可以扩充成为除环(或域)呢? 因为除环(或域)没有零因子，所以一个环 R 能被除环(或域)包含，R 必须没有零因子. 然而，对于非交换环，无零因子这个条件还不充分. 关于无零因子的非交换环可以扩充成除环的充要条件可参考 N. Jacobson: Basic Algebra Ⅰ，119(中译本，第141页). 下面我们要证明: 当 R 是交换环时，无零因子这个条件还是充分的，所用的方法类似于由整数环扩充成有理数域的方法.

定理 6.2.14 每一个无零因子交换环 R 都可以扩充为一个域 F.

定理中的无零因子交换环 R 的扩域 F 的构造为: $F=\{ab^{-1} \mid a\in R, b\in R^*\}$.

定义 6.2.8 设 R 是无零因子交换环，F 是 R 的扩域，且 $F=\{ab^{-1}\mid a\in R, b\in R^*\}$，则称 F 是 R 的**商域**(或**分式域**).

由定理及推论知，每一个无零因子交换环 R 都存在商域 F. 有理数域 **Q** 是整数环 **Z** 的商域，实数域 **R** 不是整数环 **Z** 的商域.

定理 6.2.15 设 F, F' 分别是环 R, R' 的商域，若 $R\cong R'$，则 $F\cong F'$.

在上述定理中，取 $R=R'$ 就得到下面商域的唯一性定理.

定理 6.2.16 设 F 与 F' 都是环 R 的商域，则 $F\cong F'$.

推论 6.2.17 环 R 的商域是 R 的最小扩域.

6.3 唯一分解环

在整数环 **Z** 中，每一个不等于 ± 1 的非零整数都能分解成有限个素数的乘积，而且除了因数次序以外，分解是唯一的. 在数域 F 上的一元多项式环 $F[x]$ 中，每一个次数大于等于 1 的多项式都能分解成有限个不可约多项式的乘积，而且除了因子次序和零次因式的差别以外，分解是唯一的. 一些特殊的整环也具有上面的性质.

在这一节里，我们将对一般的整环讨论元素因子分解问题，给出整环中因子分解唯一性定理成立的一些条件，并介绍几种唯一分解定理成立的整环.

6.3.1 单位、素元与可约元

定义 6.3.1 整环 I 中的可逆元 ε 称为 I 的**单位**.

ε 是 I 的单位 $\Leftrightarrow (\varepsilon)=I$. 整环 I 的全体单位关于 I 的乘法构成一个交换乘法群.

一个元数大于 2 的整环中至少有两个单位: 1 和 -1. 整数环 **Z** 只有两个单位，即 ± 1. 域 F 中的每一个非零元都是单位.

定义 6.3.2 设 $a, b\in I$，若存在 $c\in I$，使 $a=bc$，则称 b 整除 a，记作 $b\mid a$.

定义 6.3.3 设 $a, b\in I$，若 $a\mid b$ 且 $b\mid a$，则称 a 与 b 相伴，记作 $a\sim b$.

相伴关系有下列性质:

设 $a,b,c \in I$,则下列各个命题等价：

1) $a \sim b$； 2) $b = \varepsilon a$； 3) $(a) = (b)$.

相伴关系是整环 I 上的一个等价关系.

例 6.3.1 设 $I = Z[\sqrt{3}\mathrm{i}] = \{m + n\sqrt{3}\mathrm{i} \mid m,n \in \mathbf{Z}\}$.

1) 则 ε 是 I 的单位 $\Leftrightarrow |\varepsilon| = 1 \Leftrightarrow \varepsilon = \pm 1$； 2) 求 2 的相伴元.

证明 1) 若 ε 是 I 的单位,则存在 $\varepsilon' \in I$,使 $\varepsilon\varepsilon' = 1$. 两边取模的平方,得 $|\varepsilon|^2 |\varepsilon'|^2 = 1$,设 $\varepsilon = m + n\sqrt{3}\mathrm{i}$,则 $|\varepsilon|^2 = m^2 + 3n^2$,是正整数,同理,$|\varepsilon'|^2$ 也是正整数,于是 $|\varepsilon|^2 = 1$. 若 $|\varepsilon|^2 = 1$,则 $m^2 + 3n^2 = 1$,所以 $n = 0, m = \pm 1$,即 $\varepsilon = \pm 1$. 显然 ± 1 是 I 的单位.

2) 由(1)与相伴元的定义可得,2 的相伴元只有 2 与 -2.

定义 6.3.4 设 $a,b \in I$,若 $b \mid a$,但 b 不是单位,且 b 与 a 不相伴,称 b 是 a 的**真因子**.

定理 6.3.1 设 $a,b \in I$,则 b 是 a 的真因子 $\Leftrightarrow (a) \subset (b) \subset I$.

证明 因为 $b \mid a \Leftrightarrow (a) \subseteq (b)$；又 b 与 a 不相伴 $\Leftrightarrow (a) \neq (b)$；而 b 不是单位 $\Leftrightarrow (b) \neq 1$,所以命题成立.

定义 6.3.5 $0 \neq a \in I$,且 a 不是单位,若 a 在 I 中没有真因子,则称 a 是 I 的一个**不可约元**；若 a 在 I 中有真因子,则称 a 是 I 的一个**可约元**.

定理 6.3.2 设 $0 \neq a \in I$,且 a 不是单位,则 a 是可约元 $\Leftrightarrow a = bc$,且 b,c 都不是单位.

证明 设 a 是可约元,则 a 有真因子 b,于是 $a = bc$,则 c 也是 a 的真因子,从而 b 与 c 都不是单位.

反之,设 $a = bc$,b,c 都不是单位,下面证 b 是 a 的真因子,只要证 b 与 a 不相伴. 不然,若 $a \sim b$,则存在 I 的单位 ε,使 $b = \varepsilon a$,于是 $a = bc = \varepsilon ac$,由消去律得 $\varepsilon c = 1$,于是 c 是单位,矛盾,因此 a 是可约元.

注意：一个不可约元的相伴元也是不可约元.

定义 6.3.6 $0 \neq p \in I$,且 p 不是单位,若由 $p \mid ab$ 可得 $p \mid a$ 或 $p \mid b$,则称 p 是 I 的**素元**.

定理 6.3.3 在整环 I 中,每一个素元都是不可约元.

证明 设 p 是一个素元,若 $p \neq 0$,则 $p \mid ab$. 由素元定义可得 $p \mid a$ 或 $p \mid b$. 若 $p \mid a$,而由 $p = ab$ 又可得 $a \mid p$,所以 $a \sim p$,从而 b 是单位. 同理,若 $p \mid b$,则 a 是单位. 因此 p 是不可约元.

注意：逆命题在一般的整环中不成立. 看下面的例子.

例 6.3.2 设 $I = Z[\sqrt{3}\mathrm{i}] = \{m + n\sqrt{3}\mathrm{i} \mid m,n \in \mathbf{Z}\}$,则

1) I 中适合条件 $|a|^2 = 4$ 的元 a 是 I 的不可约元；

2）2 是 I 的不可约元，但不是 I 的素元.

6.3.2 唯一分解环

定义 6.3.7 设 $a,b \in I$，若存在 $d \in I$，使

1）$d \mid a, d \mid b$（这时称 d 是 a 与 b 的**公因子**）；

2）$\forall c \in I$，c 是 a 与 b 的公因子，则 $c \mid d$，称 d 是 a 与 b 的**最大公因子**.

定义 6.3.8 设 $a,b \in I$，若 a 与 b 的最大公因子是单位，则称 a 与 b **互素**.

定义 6.3.9 设 $a \in I$ 满足：

1）有一个因子分解式：$a = p_1 p_2 \cdots p_r$（p_i 是 I 中不可约元）；

2）若同时又有因子分解式：$a = q_1 q_2 \cdots q_s$（q_i 是 I 中不可约元），

那么 $s = r$，并且可以适当调换因子的次序，使 $q_i = p_i (i = 1,2,\cdots,r)$，则称 a 为 I 中的**唯一分解元**，并称 r 是 a 的长.

若整环 I 中每一个既不是零又不是单位的元都是唯一分解元，则称 I 是**唯一分解环**.

设 a 是唯一分解元，若在 a 的分解式中，有 t 个不可约因子 $p_1, p_2, \cdots, p_t$ 互不相伴，且其他的不可约因子都与某个 p_i 相伴，则 a 的分解式可以写作：$a = \varepsilon p_1^{e_1} p_2^{e_2} \cdots p_t^{e_t}$，其中 ε 是单位，$e_i \in \mathbf{N}$，这个式子称为 a 的**标准分解式**.

按此定义，整数环 $\mathbf{Z}$ 与数域 F 上一元多项式环 $F[x]$ 都是唯一分解环. 但一般的整环未必是唯一分解环.

例 6.3.3 整环 $I = \mathbf{Z}[\sqrt{3}\mathrm{i}]$ 不是唯一分解环.

引理 6.3.4 在一个唯一分解环 I 中，若元 a 的不可约因子分解已知，则可确定出 a 的所有真因子，且元 a 的长大于其任一真因子的长.

证明 设 b 是 a 的真因子，$a = bc$，b,c 都不是单位. 因为 I 是唯一分解环，所以 $a = p_1 p_2 \cdots p_r, r > 0, b = p_1' p_2' \cdots p_s', s > 0, c = p_1'' p_2'' \cdots p_t'', t > 0$，$p_i, p_i', p_i''$ 是不可约元，则由因子分解的唯一性，得 $s + t = r$，且 $p_j' \sim p_{i_j}$，当 $j \neq k$ 时 $i_j \neq i_k$，从而 $r > s, b \sim p_{i_1} p_{i_2} \cdots p_{i_s}$.

定理 6.3.5 在一个唯一分解环 I 中，任意两个元都有最大公因子.

定理 6.3.6 在一个唯一分解环 I 中，每一个不可约元都是素元.

证明 由定理 6.3.5 知 I 中任意两个元的最大公因子都存在，设 p 是 I 的一个不可约元，若 p 不是素元，则存在 $a,b \in I$，使 $p \mid ab$，而 p 不整除 a，且 p 不整除 b，于是 $(p,a) \sim 1, (p,b) \sim 1$，则 $(p,ab) \sim 1$，从而 p 不整除 ab，矛盾. 因此 p 是素元.

推论 6.3.7 若整环 I 中任两个元的最大公因子都存在，则 I 中的每个不可约元都是素元.

定理 6.3.8 若整环 I 满足：

1）I 中每一个既不是零又不是单位的元 a 都有一个因子分解：$a = p_1 p_2 \cdots p_r$（p_i

不可约)；

2) 若 I 的每一个不可约元 p 都是素元，则 I 是唯一分解环.

定理 6.3.9 若整环 I 满足：

1) I 中每一个既不是零又不是单位的元 a 都有一个因子分解：$a = p_1 p_2 \cdots p_r$；

2) I 的任意两个元都存在最大公因子，

则 I 是唯一分解环.

6.3.3 主理想环

定义 6.3.9 若整环 I 的每一个理想都是主理想，则称 I 是**主理想环**.

例 6.3.4 整数环 $\mathbf{Z}$ 是一个主理想环；域 F 上的一元多项式环 $F[x]$ 也是一个主理想环. 但是在整数环 $\mathbf{Z}$ 上的一元多项式环 $\mathbf{Z}[x]$ 中，$(2,x)$ 不是主理想，因此 $\mathbf{Z}[x]$ 不是主理想环.

引理 6.3.10 设 I 是一个主理想环，若在序列 $a_1,a_2,a_3,\cdots(a_i \in I, i=1,2,3,\cdots)$ 中每一个元都是前面一个元的真因子，则这个序列一定是有限序列.

证明 作序列的各个元生成的主理想：$(a_1),(a_2),(a_3),\cdots$

由于 a_{i+1} 是 a_i 的真因子，则 $(a_1) \subset (a_2) \subset (a_3) \subset \cdots$ 令 $T = \bigcup (a_i)$，则任意 $a,b \in T$ 及 $r \in I$，总有 $a \in (a_i), b \in (a_j)$，其中 i,j 为某两个正整数，假设 $i \leqslant j$，则 $(a_i) \subseteq (a_j)$，从而 $a \in (a_j)$，于是 $a-b, ra \in (a_j) \subseteq T$，因此 T 是 I 的一个理想. 因为 I 是主理想环，所以 $T=(d)$. 于是 $d \in T = \bigcup (a_i)$，从而 d 属于某个 (a_k). 而且，这个 a_n 一定是序列中的最后一个元，不然，在 a_n 后面还有一个元 a_{n+1}，由于 $d \in (a_n), a_{n+1} \in T=(d)$，可得 $a_n \mid d, d \mid a_{n+1}$，于是 $a_n \mid a_{n+1}$，而由假设 $a_{n+1} \mid a_n$，从而 $a_{n+1} \sim a_n$，这与 a_{n+1} 是 a_n 的真因子的假设相矛盾.

定理 6.3.11 每一个主理想环 I 都是唯一分解环.

证明 在 I 中任取一个既不是零也不是单位的元 a，假定 a 不能写成有限个不可约元的乘积，则 a 是可约元，从而有真因子 b，即存在 $c \in I$，使 $a = bc$，则 c 也是 a 的真因子，因为 a 不能写成有限个不可约元的乘积，所以 b 与 c 中至少有一个也不能写成有限个不可约元的乘积，将这个元记作 a_1. 于是 a_1 又有真因子 a_2，如此下去，得到了 I 的一个无限真因子序列：$a,a_1,a_2,\cdots$ 与引理矛盾. 因此 a 有一个不可约元的因子分解.

其次，$\forall a,b \in I$，由 a,b 生成的理想 $(a,b)=\{ar+bs \mid r,s \in I\}$.

因 I 是主理想环，所以存在 $d \in I$，使 $(a,b)=(d)$，于是 $a \in (d), b \in (d)$，从而 $d \mid a, d \mid b$，又若 $c \mid a, c \mid b$，则 $a \in (c), b \in (c)$，对任意 $r,s \in I$，有 $ar+bs \in (c)$，即 $(d) \subseteq (c)$，从而 $c \mid d$，因此 d 是 a,b 的一个最大公因子.

由于整数环 $\mathbf{Z}$ 与域 F 上的一元多项式环 $F[x]$ 都是一个主理想环，因此 $\mathbf{Z}$ 与 $F[x]$ 都是唯一分解环.

注意：该定理的逆命题不成立，即唯一分解环未必是主理想环.

例 6.3.5 如 $Z[x]$不是主理想环，但 $Z[x]$是唯一分解环.

我们知道，在一个唯一分解环 I 中，任意两个元都有最大公因子. 在一个主理想环中两个元的最大公因子还有更进一步的性质：

定理 6.3.12 设 I 是主理想环，$a,b\in I$，则$(a,b)=(d)\Leftrightarrow d$ 是 a 与 b 的一个最大公因子.

证明 *必要性* 设 $(a,b)=(d)$，则 $a\in(d)$，$b\in(d)$，于是 $d\mid a$，$d\mid b$，所以 d 是 a,b 的一个公因子，同时 $d\in(a,b)$，于是 $d=au+bv$，$u,v\in I$. 设 c 是 a,b 的任意一个公因子，即 $c\mid a$，$c\mid b$，从而由上式得 $c\mid d$. 因此 d 是 a,b 的一个最大公因子.

充分性 因 I 是主理想环，存在 $c\in I$，使$(a,b)=(c)$，从而 c 是 a 与 b 的一个最大公因子，由假设 d 是 a,b 的一个最大公因子，则 $d\sim c$，因此$(d)=(c)=(a,b)$.

该定理可以推广为：设 I 是主理想环，$a_1,a_2,\cdots,a_r\in I$，则 $(a_1,a_2,\cdots,a_r)=(d)\Leftrightarrow d$ 是 $a_1,a_2,\cdots,a_r$ 的一个最大公因子.

定理 6.3.13 设 I 是一个主理想环，p 是 I 中的非零元，则(p)是 I 的极大理想$\Leftrightarrow p$ 是 I 的不可约元.

证明 *充分性* 设 p 是 I 的不可约元. 若 I 中有一个理想 A，使$(p)\subseteq A$，则因为 I 是主理想环，所以存在 $a\in I$，使 $A=(a)$. 于是$(p)\subseteq(a)$，即 $a\mid p$，且 a 与 p 不相伴. 而 p 是不可约元，从而 a 是单位，即 $A=(a)=I$，因此(p)是 I 的极大理想.

必要性 因为(p)是 I 的极大理想，所以$(p)\neq I$，p 不是单位. 现设 $b\mid p$，则$(p)\subseteq(b)$. 因为(p)是极大理想，所以$(b)=(p)$或$(b)=I$，即 $b\sim p$ 或 b 是单位. 因此 p 是不可约元.

6.3.4 欧氏环

定义 6.3.10 设 I 是整环，若

(1)存在一个由 $I^*=I-\{0\}$ 到非负整数集 $\mathbf{N}\cup\{0\}$ 的映射 φ；

(2) 对 $\forall a\in I^*$，$b\in I$，存在 $q,r\in I$，使 $b=aq+r$，$r=0$ 或 $\varphi(r)<\varphi(a)$，

则称 I 是一个**欧氏环**.

例 6.3.6 整数环 $\mathbf{Z}$ 是一个欧氏环.

证明 首先整数环 $\mathbf{Z}$ 是一个整环. 令 $\varphi:\mathbf{Z}^*\to\mathbf{N}\cup\{0\}$，$a\to|a|$，其中 $|a|$ 是 a 的绝对值，则 φ 是一个 $\mathbf{Z}^*$ 到 $\mathbf{N}\cup\{0\}$ 的映射，并且对 $\forall a\in\mathbf{Z}^*$，$b\in\mathbf{Z}$，存在 $q,r\in I$，使 $b=aq+r$，$r=0$ 或 $\varphi(r)=|r|<|a|=\varphi(a)$. 因此 $\mathbf{Z}$ 是一个欧氏环.

例 6.3.7 Gauss 整数环 $Z[\mathrm{i}]=\{m+n\mathrm{i}\mid m,n\in\mathbf{Z}\}$ 是一个欧氏环.

证明 $Z[\mathrm{i}]$ 是复数域 $\mathbf{C}$ 的一个子环，且 $1\in Z[\mathrm{i}]$，于是 $Z[\mathrm{i}]$ 是一个整环. 令 $\varphi:Z[\mathrm{i}]^*\to\mathbf{N}\cup\{0\}$，$a\to|a|^2$，其中 $|a|$ 是 a 的模，则 φ 是一个 $Z[\mathrm{i}]2$ 到 $\mathbf{N}\cup\{0\}$ 的

映射.

下面证明 $\forall a \in Z[\mathrm{i}]^*, b \in Z[\mathrm{i}]$,存在 $q, r \in Z[\mathrm{i}]$,使 $b = aq + r, r = 0$ 或 $\varphi(r) < \varphi(a)$.

设 $a^{-1}b = u + v\mathrm{i}$,其中 $u, v \in \mathbf{Q}$,现取 u', v' 分别是与 u, v 最接近的整数,令 $k = u - u', h = v - v'$,则 $|k| = |u - u'| \leqslant 1/2$, $|h| = |v - v'| \leqslant 1/2$,于是 $b = a(u + v\mathrm{i}) = a[(u' + k) + (v' + h)\mathrm{i}] = a(u' + v'\mathrm{i}) + a(k + h\mathrm{i}) = aq + r$,其中 $q = u' + v'\mathrm{i}, r = k + h\mathrm{i} \in Z[\mathrm{i}]$,因为 $r = b - aq$,所以 $r \in Z[\mathrm{i}]$.若 $r \neq 0$,则 $\varphi(r) = |r|^2 = |a|^2 |k + h\mathrm{i}|^2 = |a|^2(k^2 + h^2) \leqslant |a|^2(1/4 + 1/4) = 1/2\,\varphi(a) < \varphi(a)$.因此 $Z[\mathrm{i}]$ 是欧氏环.

定理 6.3.14 域 F 上的一元多项式环 $F[x]$是欧氏环.

证明 显然 $F[x]$ 是一个整环.令 φ: $F[x]^* \to \mathbf{N} \cup \{0\}, f(x) \to \deg f(x)$,其中 $\deg f(z)$ 是 $f(x)$ 的次数,则 φ 是一个 $F[x]^*$ 到 $\mathbf{N} \cup \{0\}$ 的映射.对 $\forall f(x) \in F[x]^*$, $g(x) \in F[x]$,由于 $f(x)$ 的最高项系数不等于零,即是域F的单位,从而存在 $q(x), r(x) \in F[x]$,使 $g(x) = f(x)q(x) + r(x)$,其中 $r(x) = 0$ 或 $\varphi(r(x)) = \deg r(x) < \deg f(x) = \varphi(f(x))$.因此 $F[x]$ 是一个欧氏环.

定理 6.3.15 每一个欧氏环 I 是主理想环,从而是唯一分解环.

注意:定理 6.3.15 的逆命题不成立.例如,整环$\{(a + b\sqrt{19}\mathrm{i})/2 \mid a, b \in \mathbf{Z}, a \equiv b(\mathrm{mod}2)\}$是主理想环,但不是欧氏环.

6.4 唯一分解环上的一元多项式环

我们已经看到,域 F 上的一元多项式环 $F[x]$是唯一分解环,现在我们将推广这一结果,证明唯一分解环 I 上的一元多项式环 $I[x]$也是唯一分解环.

6.4.1 本原多项式

定理 6.4.1 1) $I[x]$是一个整环,I 的单位元就是 $I[x]$的单位元;

2) $I[x]$中的单位恰是 I 中的单位.

定义 6.4.1 若 $f(x) \in I[x]$ 的系数的最大公因子是单位,则称 $f(x)$ 是 $I[x]$ 的一个**本原多项式**.

定理 6.4.2 本原多项式有如下性质:

1) 本原多项式不能是零多项式;

2) 设 $f(x)$ 是零次多项式,则 $f(x)$ 为本原多项式当且仅当 $f(x)$ 为单位;

3) 与本原多项式相伴的多项式也是本原多项式;

4) 若本原多项式 $f(x)$ 可约,则 $f(x) = f_1(x)f_2(x)$,其中 $0 < \deg f_i(x) < \deg f(x)$;

5) 设 $0 \neq f(x) \in I[x]$,则 $f(x) = df_1(x)$,其中 $d \in I, f_1(x)$ 是本原多项式,且

这个分解除相差单位因子外是唯一的.

令 Q 是 I 的商域,上面的结果可以推广到 Q 上的一元多项式环 $Q[x]$) 中.

引理 6.4.3 设 Q 是 I 的商域,$0 \neq f(x) \in Q[x]$,则 $f(x)$ 可表示为 $f(x) = rf_1(x)$,其中 $0 \neq r \in Q$,$f_1(x)$ 是 $I[x]$ 中的一个本原多项式,并且这个分解除相差 I 中的单位因子外是唯一的.

证明 设 $f(x) = q_0 + q_1 x + \cdots + q_n x^n$,$q_i \in Q$,于是 $q_i = a_i/b_i$,$a_i, b_i \in I$. $b_i \neq 0$. 令 $b = b_0 b_1 \cdots b_n$,则 $0 \neq bf(x) \in I[x]$,由定理 6.4.2(5),得 $bf(x) = cf_1(x)$,其中 $c \in I$,$f_1(x) \in I[x]$ 是一个本原多项式. 于是 $f(x) = rf_1(x)$,$r = cb^{-1} \in Q$.

其次,若又有 $f(x) = sf_2(x)$,$s = ed^{-1} \in Q$,其中 $e, d \in I$,$d \neq 0$,$f_2(x) \in I[x]$ 是本原多项式. 于是 $cdf_1(x) \doteq bef_2(x) \in I[x]$,由定理 6.4.2(5) 中关于分解的唯一性,得$cd \sim be$,$f_1(x) \sim f_2(x)$,从而 $f_1(x) = \varepsilon f_2(x)$,$\varepsilon$ 是单位.

引理 6.4.4 (Gauss) 在 $I[x]$ 中,设 $f(x) = g(x)h(x)$,$f(x)$ 是本原多项式 $\Leftrightarrow g(x)$ 与 $h(x)$ 都是本原多项式.

证明 必要性显然成立,下证充分性,设

$$g(x) = a_0 + a_1 x + \cdots + a_n x^n, h(x) = b_0 + b_1 x + \cdots + b_m x^m,$$

$$f(x) = g(x)h(x) = c_0 + c_1 x + \cdots + a_n x^{n+m}, \text{其中 } c_k = \sum_{i+j=k} a_i b_j,$$

因为 $g(x)$,$h(x)$ 是本原多项式,所以 $g(x) \neq 0$,$h(x) \neq 0$,从而 $f(x) \neq 0$. 若 $f(x)$ 不是本原多项式,则 $c_0, c_1, \cdots, c_{m+n}$ 的一个最大公因子 d 不是 I 的单位. 而 $f(x) \neq 0$,于是 $d \neq 0$. 由于 I 是唯一分解环,于是存在 I 的不可约元 p,使 $p \mid d$,从而 $p \mid c_k$,$k = 0, 1, \cdots, m+n$,因为 $g(x)$,$h(x)$ 是本原的,所以 p 不能整除所有的 a_i,也不能整除所有的 b_j. 设 a_i 中不能被 p 整除的下标最小的为 a_i,b_j 中不能被 p 整除的下标最小的为 b_j,考察 $f(x)$ 中 x^{i+j} 的系数 $c_{i+j} = a_0 b_{i+j} + a_1 b_{i+j-1} + \cdots + a_i b_j + \cdots + a_{i+j} b_0$,该式中 $a_0, a_1, \cdots, a_{i-1}, b_{j-1}, \cdots, b_0, c_{i+j}$ 都能被 p 整除,从而 $p \mid a_i b_j$. 由于唯一分解环中的不可约元也是素元,从而 $p \mid a_i$ 或 $p \mid b_j$,这与 a_i, b_j 的取法矛盾,因此 $f(x)$ 是本原多项式.

定理 6.4.5 设 Q 是 I 的商域,$g(x)$ 是 $I[x]$ 中次数大于零的不可约多项式,则 $g(x)$ 在 $Q[x]$ 中也是不可约的.

证明 因为 $q(x)$ 的次数大于零,且在 $I[x]$ 中不可约,所以其系数的最大公因子只能是单位,从而 $g(x)$ 是本原多项式. 若 $q(x)$ 在 $Q[x]$ 中是可约的,则 $q_1(x)q_2(x)$,其中 $q_1(x), q_2(x) \in Q[x]$,且都不是单位,即不是 Q 中的非零元,所以 $q_1(x), q_2(x)$ 都是 $Q[x]$ 中次数大于零的多项式,于是 $q_i(x) = r_i h_i(x)$,$i = 1, 2$,其中 $r_i \in Q$,$h_i(x)$ 是 $I[x]$ 中的本原多项式,从而 $q(x) = r_1 r_2 h_1(x) h_2(x)$,其中 $r_1, r_2 \in Q$,且由引理 6.4.4 知 $h_1(x)$,$h_2(x)$ 是 $I[x]$ 中的本原多项式,得 $q(x) = \varepsilon h_1(x) h_2(x)$,其中 ε 是 I 的单位. 然而 $\deg h_i(x) = \deg q_i(x) > 0$,这与 $q(x)$ 在 $I[x]$ 中的不可约性矛盾. 因此 $q(x)$ 在 $Q[x]$ 中

不可约.

引理 6.4.6 $I[x]$中次数大于零的本原多项式 $f_0(x)$是 $I[x]$中的唯一分解元.

证明 先证 $f_0(x)$可以分解成不可约多项式的乘积,若 $f_0(x)$是不可约多项式,则结论已成立. 现设 $f_0(x)$ 可约,则 $f_0(x) = f_1(x)f_2(x)$,其中 $0 < \deg f_i(x) < \deg f_0(x)$,$i = 1,2$. 因为 $f_0(x)$ 是本原多项式,所以 $f_1(x)$,$f_2(x)$ 也是本原多项式. 这样,若 $f_1(x)$ 或 $f_2(x)$ 也是可约的,又可把它们分解成次数更低的本原多项式的乘积. 由于$f_0(x)$的次数是一个有限正整数,从而总有分解式:$f_0(x) = q_1(x)q_2(x)\cdots q_n(x)$,其中每个 $q_i(x)$ 是 $I[x]$ 中的不可约本原多项式,且次数大于零. 若 $f_0(x)$ 另有一个分解式:$f_0(x) = p_1(x)p_2(x)\cdots p_m(x)$,其中每个 $p_i(x)$ 是$I[x]$ 中的不可约本原多项式,且次数大于零,则,$q_i(x)$ 与 $p_i(x)$ 在 $Q[x]$ 中也是不可约的,从而上面 $f_0(x)$ 的两个分解式也是 $f_0(x)$ 在 $Q[x]$ 中的两个不可约元的因子分解,而 $Q[x]$ 是唯一分解环,所以 $n = m$,且适当调换因子的次序,在 $Q[x]$ 中有 $q_i(x) \sim p_i(x)$,即存在 $Q[x]$ 的单位 a_i,使 $q_i(x) = a_ip_i(x)$.

6.4.2 唯一分解环

已知 $Q[x]$ 的单位就是 I 的商域Q 的单位,从而 $0 \neq a_i \in Q$,所以 $p_i(x) = \varepsilon_i q_i(x)$,$\varepsilon_i$ 是 I 的单位,即在 $I[x]$ 中也有 $q_i(x) \sim p_i(x)$. 因此 $f_0(x)$ 是 $I[x]$ 中的唯一分解元.

定理 6.4.7 设 I 是唯一分解环,则 $I[x]$也是唯一分解环.

证明 设 $f(x) \in I[x]$,$f(x) \neq 0$,且 $f(x)$ 不是单位. 若 $f(x) \in I$,因 I 是唯一分解环,则 $f(x)$ 是唯一分解元. 若 $f(x)$ 是本原多项式,则 $f(x) \notin I$,即 $\deg f(x) > 0$,于是 $f(x)$ 也是唯一分解元. 以下设 $f(x)$ 是次数大于零的非本原多项式. 又 $f(x) = df_0(x)$,其中 $d \in I$ 不是单位,$f_0(x)$ 是次数大于零的本原多项式. 因为 I 是唯一分解环,所以d在I 中有因子分解:$d = p_1p_2\cdots p_r$,其中 p_i 是I 中不可约元. 因 $f_0(x)$ 在 $I[x]$ 中有因子分解:$f_0(x) = q_1(x)q_2(x) \cdots q_n(x)$,其中 $q_i(x)$ 是不可约本原多项式,且次数大于 0. 由上两式得到 $f(x)$ 在 $I[x]$ 中的不可约因子分解:

$$f(x) = p_1p_2\cdots p_rq_1(x)q_2(x)\cdots q_n(x).$$

假设 $f(x)$ 在 $I[x]$ 中另有一个不可约因子分解:$f(x) = p_1'p_2'\cdots p_s'q_1'(x)q_2'(x)\cdots q_m'(x)$,其中 $p_i'\in I$,$q_i'(x)$ 是不可约本原多项式,且次数大于 0. 令 $d' = p_1'p_2'\cdots p_s'$, $f_1(x) = g\, q_1'(x)q_2'(x)\cdots q'_m(x)$,则 $f(x) = df_0(x) = d'f_1(x)$,其中 $d,d' \in I$,$f_0(x)$,$f_1(x)$ 为 $I[x]$ 中本原多项式,则 $d \sim d'$,$f_0(x) \sim f_1(x)$. 由于 I 是唯一分解环,得 $d' = \varepsilon d$,$f_0(x) = \varepsilon^{-1}f_1(x)$ 都是唯一分解元,从而 $r = s$,$n = m$,且适当调换因子的次序,可使 $p_i \sim p'_i$,$q_i(x) \sim q'_i(x)$. 因此 $f(x)$ 是唯一分解元. 综上所证可得 $I[x]$ 是一个唯一分解环.

该定理可以推广为：设 I 是唯一分解环，则 $I[x_1,x_2,\cdots,x_n]$ 也是唯一分解环.

由此可见，整数环 $\mathbf{Z}$ 上的一元多项式环 $\mathbf{Z}[x]$ 是一个唯一分解环，域 F 上二元多项式环 $F[x,y]$ 也是一个唯一分解环. 但是，我们知道 $\mathbf{Z}[x]$ 不是主理想环，$F[x,y]$ 也不是主理想环，因此，唯一分解环比主理想环更为广泛.

6.4.3 多项式的根

定理 6.4.8(Eisenstein 判别法) 设 $n \geqslant 1, f(x) = a_0 + a_1x + \cdots + a_nx^n \in I[x]$. 若存在 J 中不可约元 p，使：1) p 不整除 a_n；2) $p \mid a_i (i = 0,1,\cdots,n-1)$；3) p^2 不整除 a_n，则 $f(x)$ 在 $I[x]$ 中不能分解成两个次数比 n 小的多项式乘积，即 $f(x)$ 在 $Q[x]$ 中不可约.

证明 反证法 设 $f(x) = g(x)h(x)$，其中 $g(x),h(x) \in I[x]$，且 $0 < \deg g(x)$，$\deg h(x) < n$，令

$$g(x) = b_0 + b_1x + \cdots + b_mx^m, m \geqslant 1; h(x) = c_0 + c_1x + \cdots + c_tx^t,\ t \geqslant 1;$$

因为 $p \mid a_0$，即 $p \mid b_0c_0$，又 p^2 不整除 a_0，所以可设 $p \mid b_0$，而 p 不整除 c_0. 因为 p 不整除 a_n，所以 p 不能整除 $g(x)$ 的所有系数，设第一个不被 p 整除的 $b_j (1 \leqslant j \leqslant m)$，由于 I 是唯一分解环，不可约元 p 也是素元，从而 p 不整除 b_jc_0. 于是 p 不整除 $b_jc_0 + b_{j-1}c_1 + \cdots + b_0c_j$，即 p 不整除 $a_j (j \leqslant m < n)$，这与假设矛盾，因此 $f(x)$ 在 $I[x]$ 中不能分解成两个次数比 n 小的多项式乘积，即 $f(x)$ 在 $Q[x]$ 中不可约.

例 6.4.1 $f(x,y)=4x^2+5(y-1)x+3(y-1)$ 是 $\mathbf{Z}[x,\ y]$ 中的不可约多项式.

证明 因 $f(x,\ y)$ 系数的最大公因子是单位. 又 $\mathbf{Z}[x,\ y] = \mathbf{Z}[y][x]$，所以 $f(x,\ y)$ 是系数在 $\mathbf{Z}[y]$ 中的 x 的多项式. 而 $y-1$ 是唯一分解环 $\mathbf{Z}[y]$ 中的不可约元，用 Eisenstein 判别法得 $f(x,\ y)$ 是 $\mathbf{Z}[x,\ y]$ 中不可约多项式.

最后，我们讨论整环 I 上的一元多项式环 $I[x]$ 中多项式的根的概念和性质，它与 $I[x]$ 中多项式的因子分解有关，所得结果与高等代数中讨论的数域 F 上一元多项式环 $F[x]$ 中相应的内容基本相似.

定义 6.4.2 设 $R[x]$ 是环 R 上的一元多项式环，$f(x) \in R[x]$，$a \in R$，若 $f(a) = 0$，则称 a 是 $f(x)$ 的一个**根**.

定理 6.4.9(因式定理) 设 $f(x) \in I[x]$，$a \in I$，则 a 是 $f(x)$ 的一个根 $\Leftrightarrow x-a \mid f(x)$.

证明 若 $x-a \mid f(x)$，则 $f(x) = (x-a)g(x)$，从而 $f(a) = 0$，所以 a 是 $f(x)$ 的一个根. 若 a 是 $f(x)$ 的一个根，因为 $x-a$ 的首项系数 1 是单位，则存在 $q(x) \in I[x]$，$r \in I$，使 $f(x) = (x-a)q(x) + r$，于是 $f(a) = (a-a)q(a) + r$. 因为 $f(a) = 0$，所以 $r = 0$. 从而 $f(x) = (x-a)q(x)$，即 $x-a \mid f(x)$.

定理 6.4.10 设 $f(x) \in I[x]$，$a_1, a_2, \cdots, a_n$ 是 I 中 n 个互不相同的元，则 $a_1, a_2, \cdots, a_n$ 是 $f(x)$ 的根 $\Leftrightarrow (x-a_1)(x-a_2)\cdots(x-a_n) \mid f(x)$.

证明 若 $(x-a_1)(x-a_2)\cdots(x-a_n) \mid f(x)$，则 $x-a_i \mid f(x)$，$i=1,2,\cdots,n$. 由此得 a_i 都是 $f(x)$ 的根.

反之，若 $a_1, a_2, \cdots, a_n$ 是 $f(x)$ 的根，则 $f(x)=(x-a_1)f_1(x)$，于是 $0=f(a_2)=(a_2-a_1)f_1(a_2)$，但是，$a_2-a_1\neq 0$，$I$ 没有零因子，所以 $f_1(a_2)=0$，即 a_2 是 $f_1(x)$ 的根.

又 $f_1(x)=(x-a_2)f_2(x)$，于是 $f(x)=(x-a_1)(x-a_2)f_2(x)$. 如此下去，得到：

$$f(x)=(x-a_1)(x-a_2)\cdots(x-a_n)f_n(x).$$

例 6.4.2 设 $f(x)=x^3-x$ 是 $\mathbf{Z}_3[x]$ 中一个多项式，则 $\mathbf{Z}_3$ 中的每一个元都是 $f(x)$ 的根.

证明 $\mathbf{Z}_3=\{[0],[1],[2]\}$，$f(x)=x^3-x=x(x-[1])(x-[2])$，得证.

推论 6.4.11 设 $f(x)\in I[x]$，若 $\deg f(x)=n$，则 $f(x)$ 在 I 中至多有 n 个不同的根.

例如，在 $\mathbf{Z}_6[x]$ 中，二次多项式 $f(x)=x^2-x$ 有四个不同的根：$[0],[1],[3],[4]$.

6.5 域的扩张

早在 19 世纪初，Galois 在研究代数方程解法时就提出了域的概念，后来 Dedekind 和 Kronecker 在不同背景下也提出了域的概念. 系统研究域的理论始于 Weber，在 Weber等人的影响下，Steinitz 对抽象域进行了系统研究，于 1910 年发表论文"域的代数理论"，对域论本身以及相关科学的发展产生了重大影响. 域是许多数学分支(如代数、代数数论、代数几何等)研究的基础. 由于域只有平凡理想，因此无法通过域的理想来研究域. 要研究域，必须采取别的方法，其中最基本的方法就是对域进行扩张，域的扩张起源于数域的扩张.

6.5.1 素域

定义 6.5.1 若域 F 是域 E 的一个子域，则称 E 为 F 的一个**扩域**.

任何数域都包含有理数域，即有理数域是最小的数域，它不含任何真子域.

定义 6.5.2 若域 Δ 不含真子域，则称 Δ 是一个**素域**.

这样，有理数域是一个素域. 显然以素数 p 为模的剩余类域 $\mathbf{Z}/(p)$也是素域.

定理 6.5.1 设 Δ 是一个素域，则

1) 当 $\mathrm{char}\Delta \to \infty$ 时，$\Delta \cong \mathbf{Q}$；

2) 当 $\mathrm{char}\Delta = p$ 时，$\Delta \cong \mathbf{Z}/(p)$，其中 p 是素数.

在同构意义下这就是全部的素域.

证明　域 Δ 包含有单位元 e，则 Δ 也包含所有的 $ne(n \in \mathbf{Z})$，取 $S = \{ne \mid n \in \mathbf{Z}\}$，令 $\varphi: n \to ne$ 是整数环 $\mathbf{Z}$ 到 S 的一个同态映射.

1）当 $\mathrm{char}\Delta \to \infty$ 时，φ 是同构映射，从而 $Z \cong S$，但 Δ 包含 S 的商域，因同构的环的商域也同构，$\mathbf{Z}$ 的商域是有理数域 $\mathbf{Q}$，又由于 Δ 是素域，所以 S 在 Δ 中的商域就是 Δ 本身，从而 $\Delta \cong \mathbf{Q}$.

2）当 $\mathrm{char}\Delta = p$ 时，由于 $\varphi: p \to pe = 0$，即 $p \in \mathrm{Ker}\varphi$，所以 $(p) \subseteq \mathrm{Ker}\varphi$，又由于 (p) 是 $\mathbf{Z}$ 的最大理想，另一方面，$1 \to e \neq 0$，所以 $\mathrm{Ker}\varphi \neq \mathbf{Z}$，有 $\mathrm{Ker}\varphi = (p)$，则 $\mathbf{Z}/(p) \cong S$.

由于 $\mathbf{Z}/(p)$ 是域，故 S 是域. 但 Δ 是素域，故 $S = \Delta$，从而有 $\Delta \cong Z/(p)$，这样，在同构意义下，素域只有 $\mathbf{Q}$ 以及模 p(素数) 剩余类域.

推论 6.5.2　设 E 是一个域，如果 E 的特征是∞，则 E 包含一个与 $\mathbf{Q}$ 同构的素域；如果 E 的特征是素数 p，则 E 包含一个与 $\mathbf{Z}/(p)$同构的素域.

推论 6.5.3　每个域都包含一个素域且只包含一个素域.

6.5.2　扩域

下面粗略考察一下 $F(S)$中的元素是些什么样子.

在 S 中任取有限个元素 $\alpha_1, \alpha_2, \cdots, \alpha_n$，令 $f(\alpha_1, \alpha_2, \cdots, \alpha_n)$ 为系数属于 F 的关于 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 的任意一个多项式，它是 $F(S)$ 中一个确定的元素. 由于 $F(S)$ 是一个域，因此 $F(S)$ 也包含这样两个多项式的商 $f_1(\alpha_1, \alpha_2, \cdots, \alpha_n)/f_2(\alpha_1, \alpha_2, \cdots, \alpha_n)$，其中 $f_2(\alpha_1, \alpha_2, \cdots, \alpha_n) \neq 0$，另一方面，如果让 $\alpha_1, \alpha_2, \cdots, \alpha_n$ 在 S 中任意变动，n 也不固定，那么一切这样的有理分式显然作成一个包含 FUS 的子域，因此

$$F\{S\} = \{f_1(\alpha_1, \alpha_2, \cdots, \alpha_n)/f_2(\alpha_1, \alpha_2, \cdots, \alpha_n) \mid \alpha_1, \alpha_2, \cdots, \alpha_n \in S, n = 1, 2, \cdots\}$$

在 E 中适当选择 S 就能使 $E = F(5)$.

定理 6.5.4　令 E 是域 F 的一个扩域，而 S_1, S_2 是 E 的两个子集，则

$$F(S_1)(S_2) = F(S_2)(S_1) = F(S_1 \cup S_2).$$

对于域 $F(\alpha_1, \alpha_2, \cdots, \alpha_n)$ 有

$$F(\alpha_1, \alpha_2, \cdots, \alpha_n) = F(\alpha_1)(\alpha_2)\cdots(\alpha_n).$$

这就是说，讨论扩域 $F(\alpha_1, \alpha_2, \cdots, \alpha_n)$ 又可归结为讨论向 F 添加一个元素而得到的扩域，这种扩域很重要.

定义 6.5.3　设 E 是域 F 的一个扩域，$\alpha \in E$，如果存在 F 上非零多项式 $f(x)$使 $f(\alpha)=0$，则称 α 为 F 上的一个**代数元**. 否则，称 α 为 F 上的一个**超越元**.

例 6.5.1　$\sqrt{2}, \sqrt{3}, \mathrm{i}$ 都是有理数域 $\mathbf{Q}$ 上的代数元，而圆周率 π 则是 $\mathbf{Q}$ 上的一个超越

元.但是应注意,π 是实数域上的一个代数元.

$\mathbf{Q}(\sqrt{2})$ 是有理数域 $\mathbf{Q}$ 的一个单代数扩域;而 $\mathbf{Q}(\pi)$ 则是 $\mathbf{Q}$ 的一个单超越扩域.

定义 6.5.4 设 α 是域 F 上的一个代数元,则 F 上首系数为 1 且以 α 为根的次数最低的多项式是存在的,称为 α 在 F 上的**最小多项式**.

若 α 的最小多项式的次数是 n,则称 α 是 F 上的一个 **n 次代数元**.

例 6.5.2 i 在 $\mathbf{Q}$ 上的最小多项式是 x^2+1,因此 i 是 $\mathbf{Q}$ 上的一个 2 次代数元;$\sqrt[3]{2}$ 在 $\mathbf{Q}$ 上的最小多项式是 x^3-2,因此 $\sqrt[3]{2}$ 是 $\mathbf{Q}$ 上的一个 3 次代数元.

定理 6.5.5 域 F 上的代数元 α 在 F 上的最小多项式是唯一的,α 在 F 上的最小多项式 $p(x)$ 在 F 上不可约,又若 $f(x)$ 是一个多项式且 $f(\alpha)=0$,则 $p(x)\mid f(x)$.

定理 6.5.6 设 $F[x]$ 为域 F 上未定元 x 的多项式环,$F(x)$ 为其分式域,则

1) 当 α 为 F 上的超越元时,$F(\alpha)\cong F(x)$;

2) 当 α 为 F 上的代数元时,$F(\alpha)\cong F[x]/<p(x)>$,其中 $p(x)$ 为 α 在 F 上的最小多项式.

证明 令 $F(\alpha)=\{f(\alpha)\mid f(x)\in F[x]\}$,则 $F[\alpha]$ 是域 $F(\alpha)$ 的一个子环,又易知 φ: $f(x)\to f(\alpha)$ 是 $F[x]$ 到 $F[\alpha]$ 的一个同态满射.

1) 当 α 为 F 上的超越元时,φ 是同构映射,于是 $F[x]\cong F[\alpha]$.又同构的环其分式域也同构,$F[x]$ 的分式域是 $F(x)$,$F[\alpha]$ 的分式域是 $F(\alpha)$,因此 $F(\alpha)\cong F(x)$;

2) 当 α 为 F 上的代数元时,设 $p(x)$ 为 α 在 F 上的最小多项式,则易知 $\mathrm{Ker}\varphi=<p(x)>$,于是由环同态基本定理,得 $F[x]/<p(x)>\cong F[\alpha]$,由于 $p(x)$ 在域 F 上不可约,$<p(x)>$ 是 $F(x)$ 的极大理想,故 $F[x]/<p(x)>$ 为域,从而 $F[\alpha]$ 也是域,但 $F(\alpha)$ 是包含 F,α 的最小域,故 $F[\alpha]=F(\alpha)$,因此,$F(\alpha)\cong F[x]/<p(x)>$.

由此定理知,当 α 是域 F 上的代数元时,有 $F(\alpha)=F[\alpha]$,即每个关于 α 的有理分式都与 α 的一个多项式(系数属于 F)相等.

6.5.3 代数扩域和超越扩域

定理 6.5.7 若 α 是域 F 上的 n 次代数元,则 $F(\alpha)$ 中的每个元素都可唯一地由 $1,\alpha,\cdots,\alpha^{n-1}$ 线性表出,即可唯一地表示成 $a_0+a_1\alpha+\cdots+a_{n-1}\alpha^{n-1}$,$a_i\in F$,也就是说,$F$ 的单代数扩域 $F(\alpha)$ 也是 F 上的一个 n 维空间,而且 $1,\alpha,\cdots,\alpha^{n-1}$ 是它的一个基.

证明 设 $p(x)$ 是 α 在 F 上的最小多项式,次数为 n,任取 $\beta\in F(\alpha)$,故可令 $\beta=f(\alpha)$,其中 $f(x)\in F[x]$,设

$$f(x)=p(x)g(x)+r(x),g(x),r(x)\in F[x],r(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1},$$

用 α 代入上式,由于 $p(\alpha)=0$,故可得

$$f(\alpha)=r(\alpha)=a_0+a_1\alpha+\cdots+a_{n-1}\alpha^{n-1},$$

即 $F(\alpha)$ 中每个元素都可用 $1,\alpha,\cdots,\alpha^{n-1}$ 线性表示.

若有 $a_0+a_1\alpha+\cdots+a_{n-1}\alpha^{n-1}=b_0+b_1\alpha+\cdots+b_{n-1}\alpha^{n-1}$,其中 $a_i,b_i\in F$,则有

$$(a_0-b_0)+(a_1-b_1)\alpha+\cdots+(a_{n-1}-b_{n-1})\alpha^{n-1},$$

即 α 是 F 上多项式 $g(x)=(a_0-b_0)+(a_1-b_1)x+\cdots+(a_{n-1}-b_{n-1})x^{n-1}$ 的根.但由于 α 是 F 上的 n 次代数元,故只有 $g(x)=0$,即 $a_i=b_i,i=0,1,\cdots,n-1$,亦即 $F(\alpha)$ 中每个元素的表示法是唯一的.

给定域 F 后,F 上的单超越扩域的存在是显然的,因为当 α 为 F 上的未定元时,$F(\alpha)$ 就是 F 的一个单超越扩域.下面证明,单代数扩域也是存在的.

定理 6.5.8 设 F 是一个域,$p(x)$ 是 F 上任意一个给定的首系数为 1 的不可约多项式,则存在 F 上单代数扩域 $F(\alpha)$,其中 α 在 F 上的最小多项式是 $p(x)$.

由上面看到,单代数扩域与单超越扩域的结构是很不相同的,一般来说,设 E 是域 F 的一个扩域,则 E 中的元素有些可能是 F 上的代数元,而另一些则可能是 F 上的超越元.

例 6.5.3 实数域 $\mathbf{R}$ 是有理数域 $\mathbf{Q}$ 的一个扩域,实的代数数都是 $\mathbf{Q}$ 上的代数元,其余的实数,例如 π、e 以及 π、e 与任何非零有理数的乘积,等等,都是 $\mathbf{Q}$ 上的超越元.

定义 6.5.5 设 E 是域 F 的一个扩域,如果 E 中每个元素都是 F 上的代数元,则称 E 是 F 的一个**代数扩域**(张).否则,称 E 是 F 的一个**超越扩域**(张).

F 中的元当然都是 F 上的代数元.如果 E 是 F 的超越扩域,又 E 中除 F 的元外,都是 F 上的超越元,则称 E 是 F 的一个**纯超越扩域**(张).

例 6.5.4 $\mathbf{R}$ 是 $\mathbf{Q}$ 的超越扩域,但不是 $\mathbf{Q}$ 的纯超越扩域;又易知域 F 上关于未定元 x 的有理分式域 $F(x)$ 是 F 的一个纯超越扩域.

6.5.4 扩域次数

下面主要讨论代数扩域,我们要问:当集合 S 中的元素都是 F 上的代数元时,$F(S)$ 中的元素是否都是 F 上的代数元?也就是说,$F(S)$ 是否为 F 的代数扩域?这个看起来似乎正确的结论,实际上并不明显,这是因为,域 $F(S)$ 除了包含 F 及 S 的全部元素外,还要包含由 F 及 S 中的元素通过加、减、乘、除所得到的一切元素,这就涉及到 F 上代数元的和、差、积、商是否仍为 F 上代数元的问题.但是,这个问题虽然是正确的,却并不明显,因而是需要证明的.为了讨论这些问题,我们先介绍扩域次数的概念.设 E 是域 F 的一个扩域,则对 E 中加法与乘法来讲,E 作成 F 上的一个向量空间,于是 E 在 F 上的维数可能有限,也可能无限.

定义 6.5.6 设 E 是 F 的一个扩域,则 E 作为 F 上向量空间的维数,叫做 E 在 F 上的**次数**,记为$[E:F]$.当$[E:F]$有限时,称 E 为 F 的**有限次扩域**,否则称 E 为 F 的**无**

限次扩域.

例 6.5.5 当 α 是域 F 上的 n 次代数元时，则 $E=F(\alpha)$ 是 F 的 n 次扩域. 特别地，复数域是实数域上的 2 次扩域，$\mathbf{Q}(i)$ 是有理数域 $\mathbf{Q}$ 上的 2 次扩域，而 $\mathbf{Q}(\pi)$ 则是 $\mathbf{Q}$ 的一个无限次扩域. 从而实数域是有理数域的一个无限次扩域.

定理 6.5.9 令 E 是 K 的扩域，K 又是 F 的扩域，则

$$[E:F]=[E:K][K:F].$$

用数学归纳法可进一步证明：

推论 6.5.10 设 $F,F_1,\cdots,F_n$ 都是域，其中后一个是前一个的子域，则

$$[F_n:F]=[F_n:F_{n-1}]\cdots\ [F_2:F_1][F_1:F].$$

例 6.5.6 由于 $Q\subseteq Q(\sqrt{2}))\subseteq Q(\sqrt{2},i)$，从而有

$$[\mathbf{Q}(\sqrt{2},i):\mathbf{Q}]=[\mathbf{Q}(\sqrt{2},i):\mathbf{Q}(\sqrt{2})]\cdot[\mathbf{Q}(\sqrt{2}):Q]=2\times 2=4.$$

定理 6.5.11 有限次扩域必是代数扩域.

推论 6.5.12 若 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 都是域 F 上的代数元，则扩域 $F(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 是 F 的有限次扩域，从而为代数扩域.

推论 6.5.13 域 F 上代数元的和、差、积、商均仍为 F 上的代数元.

定理 6.5.14 设 E 是域 F 的一个扩域，S 是 E 的一个非空子集. 如果 S 中的每个元素都是 F 上的代数元，则 $F(S)$ 是域 F 的代数扩域.

证明 任取 $\beta\in F(S)$，令

$$\beta=f_1(\alpha_1,\alpha_2,\cdots,\alpha_n)/f_2(\alpha_1,\alpha_2,\cdots,\alpha_n),$$

其中 $\alpha_1,\alpha_2,\cdots,\alpha_n\in S$，$f_1,f_2(\neq 0)$ 的系数属于 F，则 $\beta\in F(\alpha_1,\alpha_2,\cdots,\alpha_n)$. 由于 $\alpha_1,\alpha_2,\cdots,\alpha_n$ 都是域 F 上的代数元，则 $F(\alpha_1,\alpha_2,\cdots,\alpha_n)$ 是 F 的代数扩域，从而 β 是 F 上的代数元，因此 $F(S)$ 是 F 的代数扩域.

定理中的 S 可以是有限集，也可以是无限集. 当 S 是有限集时 $F(S)$ 是 F 上的有限次扩域，当然是代数扩域. 但当 S 为无限集时，$F(S)$ 是 F 的代数扩域，却不一定是 F 的有限次扩域.

定理 6.5.15 设 F,K 是域 E 的两个子域，且 $F\subseteq K$，若 E 是 K 的代数扩域，K 是 F 的代数扩域，则 E 是 F 的代数扩域.

推论 6.5.16 设 E 是域 F 的超越扩域，则在 E 中存在子域 K 满足 $F\subseteq K\subset E$，其中 K 是 F 的代数扩域，而 E 是 K 的纯超越扩域.

第7章 模 理 论

模的概念是向量空间概念的直接推广，以任意环代替向量空间的基域就可以得到这样的推广，数学内在的逻辑结构使这种推广成为“自然”. 域上有限维向量空间里线性变换的研究提供了上述思想的一个具体方法，它是线性代数的一个中心问题. 我们可以看到，用 F 上向量空间 V 里给定的线性变换 σ 把 V 转化为以 x 为不定元的多项式环 $F[x]$上的模，对这种模的深入研究，将会解决线性变换中矩阵的标准型理论和矩阵相似性问题.

模在代数学里第一次成为一个重要工具是在 19 世纪 20 年代的后期，很大程度上归功于 Emmy Noether 的工作，她最早发现了模的概念的重要作用，它是沟通代数学里两个各自独立进行的重要理论 Frobenius Burnside 和 Schur 的有限群的矩阵表示论以及 Molien，Cartan 和 Wedderburn 的代数结构论的桥梁.

本章主要对主理想整环上的有限生成模进行研究，分别给出有限生成的 Abel 群的结构理论和线性变换标准化理论.

7.1 模的定义和基本性质

7.1.1 Abel 群的自同态环

令 M 为一个交换群，在 M 里用加号“+”表示给定的二元合成，0 表示单位元，$-a$ 表示 a 的逆元以及 $ma(m \in \mathbf{Z})$ 表示 m 次幂，End$M$ 表示 M 的自同态的集合. 设 η 是从 M 映入到 M 的映射，满足 $\eta(x+y)=\eta(x)+\eta(y)\eta(0)=0$.

例 7.1.1 设 M 是一个无限循环群$(\mathbf{Z},+,0)$，那么 1 是一个生成元，如果 $\eta(1)=m$，则 $\eta(1)=\eta(x1)=x\eta(1)=xm$，因此 η 是映射 $x \mapsto mx$，$x \in M$，这里 $m=\eta(1)$. 若 m 是 Z 的任意元，由幂规律 $m(x+y)=m(x)+m(y)$，可知映射 $x \mapsto mx$ 是一个自同态. 显然映射 $x \mapsto mx$ 是把 1 映入 m. 既然自同态是被它们在生成元 1 上的作用下所确定的，显然在集合 End$M$ 与 Z 之间有一一对应，有 $\eta \Leftrightarrow \eta(1)$，它将 $\eta \in$ End$M$ 与$\eta(1)=m \in M$ 相对应.

例 7.1.2 设 m 是一个有限循环群，取 $M=(\mathbf{Z}/(n),+,[0])$，这里 n 是一个正数，

一般地 $\overline{x}$ 是陪集 $x+(n)$，那么 T 是一个生成元，而在 End $\mathbf{Z}/(n)$ 与 $\mathbf{Z}/(n)$ 之间有一一对应，它将 $\eta\in \mathrm{End}\mathbf{Z}/(n)$ 与 $\eta(T)$ 相对应.

我们可以把任何 Abel 群 M 的 End$M$ 组成一个环，如果 $\eta,\xi\in \mathrm{End}M$，那么合成 $\eta\xi\in\mathrm{End}M$，而且是可结合的，即 $(\eta\xi)\zeta=\eta(\xi)\zeta$，并且恒等映射 $1: x\mapsto x$ 是一个幺半群. 即使 M 非交换，所有这些也是成立的. 但是在可交换的情况下，这个 End M 可以构成一个环.

如果 $\eta,\xi\in\mathrm{End}M$，用 $(\eta+\xi)(x)=\eta(x)+\xi(x)$ 来定义 $\eta+\xi$，这个映射是自同态，因为

$$(\eta+\xi)(x+y)=\eta(x+y)+\xi(x+y)=(\eta)(x)+(\eta)(y)+(\xi)(x)+(\xi)(y)$$
$$=(\eta)(x)+(\xi)(x)+(\eta)(y)+(\xi)(y)=(\eta+\xi)(x)+(\eta+\xi)(y),$$

我们指出，在前面运算过程中，从第二步到第三步用到加法的可换性；其次，我们定义映射 $x\mapsto 0, x\in M$ 为 0 映射. 显然，这是一个自同态，而且对 $\forall\,\eta\in\mathrm{End}M,\eta+0=\eta=0+\eta$. 令 $-\eta$ 为映射 $x\mapsto-\eta x$，那么 $-\eta$ 是 η 和自同构 $x\mapsto -x$ 的合成. 因为 M 是交换的，所以 $-\eta\in\mathrm{End}M$，显然有 $\eta+(-\eta)=0=(-\eta)+\eta$，从

$$((\eta+\xi)+\rho)(x)=(\eta+\xi)(x)+\rho(x)=\eta(x)+\xi(x)+\rho(x),$$
$$(\eta+(\xi+\rho))(x)=\eta(x)+(\xi+\rho)(x)=(\eta)(x)+(\xi)(x)+(\rho)(x),$$

可知，加法合成适合结合律. 因为

$$(\eta+\xi)(x)=(\eta)(x)+(\xi)(x)=(\xi)(x)+(\eta)(x)=(\xi+\eta)(x),$$

所以加法律也适合交换律. 于是我们验证了(End M, +, 0) 是一个 Abel 群.

我们知道，(End$M$, ·, 1) 是一个幺半群. 现在对 $\eta,\xi,\rho\in\mathrm{End}M$ 有

$$(\rho(\eta+\xi))(x)=\rho(\eta(x)+\xi(x))=(\rho\eta)(x)+(\rho\xi)(x)=(\rho\eta+\rho\xi)(x)$$
$$((\eta+\xi)\rho)(x)=\eta(\rho(x))+\xi(\rho(x)).$$

因此在 End$M$ 中两个分配律都成立.

定义 7.1.1 设 M 是一个 Abel 群，End$M$ 表示 M 的自同态的集. 对 $\eta,\xi\in\mathrm{End}M$ 用 $(\eta\xi)(x)=\eta(\xi)(x)$，$(\eta+\xi)(x)=(\eta)(x)+(\xi)(x)$ 来定义 $\eta\xi$ 和 $\eta+\xi$，用 $1x=x$ 和 $0x=0$ 来定义 1 和 0，那么(End$M$, +, ·, 0, 1) 是一个环，并称(End$M$, +, ·, 0, 1) 或简称 End$M$ 为 Abel 群的**自同态环**.

下面我们证明类似子群的凯莱定理的环的定理.

定理 7.1.1 任一环都与一个 Abel 群的一个自同态环同构.

证明 证明的思路与证明凯莱定理的思路是一致的. 给定环 $R, M\subset R$，取 R 的加法群 $M=(R,+,0)$ 以及对 $\forall\,a$，称映射 $a_L: x\mapsto ax$ 为由 a 所决定的左乘法变换

$$a_L(x+y)=a(x+y)=ax+ay=a_Lx+a_Ly, a_L\in \mathrm{End}M.$$

并且 $$(a+b)_Lx=(a+b)x=ax+bx=a_Lx+b_Lx=(a_L+b_L)x,$$

根据自同态的定义以及$(ab)_Lx=(ab)x=a(bx)=a(b_Lx)=(a_Lb_L)x$，$1x=x$，所以 $a\mapsto a_L$ 是从环R到 $\mathrm{End}M$ 的一个同态. 由 $a_L=b_L$ 可得$a=a_L1=b_L1=b$，所以 $a\mapsto a_L$ 是一个单同态，它的象是 $\mathrm{End}M$ 的子环R_L，于是我们有 $R\cong R_L$.

研究环的右乘法也是有意义的. 定义 $a_R: x\mapsto xa$，且由$(x+y)a=xa+ya$ 可知，a_R 是$M=(a_R,+,0)$的一个自同态，从而有 $a_L: a\mapsto a_R$ 是从R到 $\mathrm{End}M$ 的一个反同态. 象 $R_R=\{a_R\}$ 是 $\mathrm{End}M$ 的一个子环R到R_R的反同构.

我们指出在 $\mathrm{End}M$ 里R_L与R_R互为中心化子，即

定理 7.1.2 在 $\mathrm{End}M$ 里 $R_L=C(R_R)$ 和 $R_L=C(R_L)$.

证明 由$(ax)b=a(xb)$，对任何$a, b\in R$，可以得到$a_Lb_R=b_Ra_L$. 设η是M的一个自同态，使得对任意$a\in R$，$a_L\eta=\eta a_L$，那么$\eta(x)=\eta(x1)=\eta(x_L1)=x_L(\eta(1))=x\eta(1)$. 因此 $\eta=\eta(1)_R\in R_R$，于是 $C(R_R)=R_L$，根据对称性有 $C(R_L)=R_L$.

7.1.2 模的定义和基本性质

模和向量空间的关系非常简单：向量空间是域上的模.

定义 7.1.2 设R是有单位元1的环，M是加群，我们称M是一个**左R模**，若存在一个$R\times M$到M的映射$(r,m)\mapsto rm$，满足下列性质：

1) 对 $\forall m\in M$ 有 $1m=m$；

2) 对 $\forall r\in R, m,n\in M$ 有 $r(m+n)=rm+rn$；

3) 对 $\forall r,s\in R, m\in M$ 有$(r+s)m=rm+sm$；

4) 对 $\forall r,s\in R, m\in M$ 有 $r(sm)=(rs)m$.

如果 $R=\mathbf{Z}$，则上述条件自然都满足，此时$rm=m+\cdots+m$，此为r个m的和，因而任何一个加群可看作一个左$\mathbf{Z}$模. 设S是环，R是S的有单位元的子环，则关于S的乘法，S是一个左R模.

特别地，环R本身可看成一个左R模. 若J是R的左理想，则关于R的乘法，J是一个左R模. 若$R=F$是域，则一个F模恰好与F上的向量空间是相同的，因而模的概念是域上向量空间概念的自然推广域.

可以类似定义右R模的概念，令τ: $(m,r)\mapsto mr$，$r\in R$为$M\times R$到M的一个映射，满足类似于左R模的那些条件. 若R为交换环，则从每一个左R模通过规定运算$mr=rm$，$m\in M$可以给出一个右R模.

定义 7.1.3 设S也是一个有单位元的环，加群M称为一个(R,S)**双模**，如果M既是左R模，又是右S模，且对任意的$r\in R, m\in M, s\in S$都有$r(ms)=(rm)s$.

由定义知,任意左 R 模都是$(R,\mathbf{Z})$ 双模,任意右 R 模都是$(\mathbf{Z},R)$ 双模,环 R 本身是(R,R) 双模. 若 R 为交换环,则任意 R 模都是(R,R) 双模.

模的定义要求数环 R 是可交换的. 非交换环上的模可能有基数不同的基. 即使有了交换性这一条件,模也与向量空间的性质完全不同. 例如,不含有任何线性无关元的模是存在的,当然,这种模没有基.

定义 7.1.4 设 S 是 R 模 M 的非空子集,令 $RS = \left\{ \sum r_i s_i \mid r_i \in R, s_i \in S, 1 \leqslant i \leqslant k, k \in \mathbf{Z}^+ \right\}$. 我们称 M 由 S 生成(S 生成 M).

若 $M = RS$, 称 S 为 M 的**生成元集**. M 中元素是 S 中元素 R 的线性组合.

R 模 M 称为有限生成的,若 S 有限,且 $M = RS$. 换句话说,M 的每个元素可写为 S 中元素 R 的线性组合.

若 $\forall\, T \subseteq S$, T 不能生成 M,则 S 称为 M 的一个**极小生成元集**.

然而,与域上有限维向量空间的任何两个基含有相同的元素个数的结论截然不同,R 模 M 的两个极小生成元集可能含有不同个数的元.

例 7.1.3 令 $M = Z_6$,则 M 可视为 Z 模. $\{1\}$ 是 M 的一个极小生成元集,而$\{2,3\}$ 也是 M 的一个极小生成元集.

定义 7.1.5 若 M 的生成元集 S 只有一个元素 x,则称 M 为循环模,记为 Rx,有时也记作 $R(x)$. 由定义得,$Rx = \{rx \mid r \in R\}$.

例 7.1.4 任何 Abel 群(记为加法的)是一个 $\mathbf{Z}$-模. 对 $a \in \mathbf{Z}$, $x \in M$ 按照通常的方法定义 ax. 从 Abel 群的性质可知,模的条件(1) ~ (4) 成立的 Abel 群就是 Z-模,这一事实容许我们把 Abel 群的理论纳入模的理论中.

例 7.1.5 我们考虑任何环 R 而取 M 为 R 的加法群$(R,+,0)$. 假定用左乘使 R 作用在 M 上,(1) ~ (4) 是显然的,所以 M 是一个左 R-模. 同样如果我们对 $x \in M$, $a \in R$,定义 xa 为 R 上的积,那么 M 是一个右 R-模.

7.1.3 子模

定义 7.1.6 设 M 是一个 R 模,N 是 M 的子群,N 称为 M 的一个 ***R* 子模**(或子模),若对任意的 $r \in R, n \in N$ 都有 $rn \in N$. 一个左 R 模的子模恰是 R 的左理想.

每一个 R 模都至少有两个子模,即它自身和零子模$\{0\}$,零子模简记为 0. 若 R 模 M 只有这两个子模,则称 M 为**单模**(或**不可约模**,或**既约模**).

一般地,零模不看作是单模,单模一定是循环模,反之不然. 例如,若 M 是循环 R 模,设为 Rx, $J \neq R$ 是 R 的非零左理想,则 Jx 是 M 的非零真子模.

定义 7.1.7 设 R 是环,M 是 R 模,令 $\mathrm{Ann}_R(M) = \{r \in R \mid r\,m \in M\}$, 则 $\mathrm{Ann}_R(M)$ 是 R 的理想,称为模 M 的**零化理想**. 如果 $\mathrm{Ann}_R(M) = 0$,则称 M 为**忠实 *R* 模**.

显然,若 R 是单环,则 M 是忠实模.

定义 7.1.8　设 N 是 M 的子模,则 N 也是 M 的子加群,而 M/N 是交换商群. 对任意的 $r\in R, m+N\in M/N$,规定 $r(m+N)=rm+N$,则易验证 M/N 是一个 R 模,称为 ***R* 商模**,简称为**商模**.

7.1.4　模的同态与同构

定义 7.1.9　设 M 和 N 是 R 模,令 $\varphi: M\to N$ 是群同态. 如果对 $\forall r\in R, m\in M$ 都有 $\varphi(rm)=r\varphi(m)$,则称 φ 是 ***R* 模同态**.

若两个模 M 和 N 是同构的,记为 $M\cong N$.

令 $\mathrm{Ker}\varphi=\{x\in M \mid \varphi(x)=0\in N\}$;$\varphi(M)=\{\varphi(m)+m\in M\}$(有时也记为 $\mathrm{Im}T$),则 $\mathrm{Ker}\varphi$ 是 M 的子模,称为 φ 的**同态核**. $\varphi(M)$ 是 N 的子模,称为 φ 的**同态像**.

类似于群,对模同态与模同构来说,我们不加证明地叙述以下重要定理:

定理 7.1.3(模同态基本定理)　设 M 和 N 是两个 R 模,$\varphi: M\to N$ 是 R 模同态,则存在同构映射 $\psi: M/K\to\varphi(M)$ 满足 $\varphi=\psi\cdot\eta$,其中 $K=\mathrm{Ker}\varphi$,$\eta: M\to M/K$ 是自然同态,而且映射 ψ 是唯一确定的.

定理 7.1.4 (第一同构定理)　设 M 是 R 模,N_1 和 N_2 是 M 的子模,则

$$(N_1+N_2)/N_1\cong N_2/(N_1\cap N_2).$$

定理 7.1.5(对应定理)　设 M,N 是 R 模,$\varphi: M\to N$ 是同态,其核为 K,则在 M 的包含 K 的子模组成的集合与 N 的子模组成的集合之间存在一一对应. 若 S 是 M 的包含 K 的子模,则 S 与 $\varphi(S)$ 对应;若 T 是 N 的子模,则 M 的子模

$$\varphi^{-1}(T)=\{m\in M \mid \varphi(m)\in T\}$$

与 T 对应;若 M_1,M_2 是 M 的包含 K 的子模,则 $M_2\subseteq M_1$ 当且仅当 $\varphi(M_2)\subseteq\varphi(M_1)$,且 $M_1/M_2\cong\varphi(M_1)/\varphi(M_2)$,其同构映射为 $x+M_2\mapsto\varphi(x)+\varphi(M_2)$.

定理 7.1.6(第二同构定理)　设 N,S 是 M 的子模,$S\subseteq N$,则

$$M/N\cong(M/S)/(N/S).$$

定理 7.1.7(Schur 引理)　任何两个单模之间的非零同态必为同构.

证明　设 M 和 N 是单模,$\varphi: M\to N$ 是 R 模同态,因为 $\mathrm{Ker}\varphi$ 是 M 的子模,因而 $\mathrm{Ker}\varphi=M$ 或 $\mathrm{Ker}\varphi=0$. 若 $\mathrm{Ker}\varphi=M$,则 $\varphi=0$,矛盾. 故 $\mathrm{Ker}\varphi=0$,即 φ 是单同态. 类似地,$\mathrm{Im}\varphi$ 是 N 的子模,从而 $\mathrm{Im}\varphi=0$ 或 $\mathrm{Im}\varphi=N$. 若 $\mathrm{Im}\varphi=0$,则 $\varphi=0$,因而必有 $\mathrm{Im}\varphi=N$,即 φ 是满同态,从而 φ 是同构.

对于任意两个子模之间的同态,我们有下列基本性质:

1) 映射 $\varphi: M\to N$ 是 R 模同态 $\Leftrightarrow$ 对 $\forall x,y\in M, r\in R, \varphi(x+y)=r\varphi(x)+r\varphi(y)$.

2) 设 $\mathrm{Hom}_R(M,N)$ 表示由 M 到 N 的所有模同态组成的集合，对 $\forall \varphi,\eta \in \mathrm{Hom}_R(M,N)$，$m \in M$，定义 $(\varphi+\eta)(m) = \varphi(m) + \eta(m)$，则 $\mathrm{Hom}_R(M,N)$ 关于这个加法是一个交换群. 设 S 是一个有单位元的环，令 M 是一个 (R,S) 双模，对于 $s \in S$，$\varphi \in \mathrm{Hom}_R(M,N)$，$m \in M$，定义 $(s\varphi)(m) = \varphi(ms)$，则 $s\varphi \in \mathrm{Hom}_R(M,N)$ 且 $\mathrm{Hom}_R(M,N)$ 是一个左 S 模.

特别地，若 $S = R$，或 R 是交换环，则 $\mathrm{Hom}_R(M,N)$ 是一个左 R 模.

3) 若 $\varphi \in \mathrm{Hom}_R(K,M)$，$\psi \in \mathrm{Hom}_R(M,N)$，则 $\psi \cdot \varphi \in \mathrm{Hom}_R(K,N)$.

4) $\mathrm{Hom}_R(M, M)$ 是一个有单位元的环，称为模 M 的自同态环，记为 $\mathrm{End}_R(M)$.

若 F 是域，U 和 V 是 F 上的向量空间，或称为 F 向量空间，则按照上面命题中的定义，$\mathrm{Hom}_F(U,V)$ 也是一个 F 向量空间.

下面令 $a \in F$，$\varphi \in \mathrm{Hom}_F(U,V)$，$u \in U$，定义 $(a\varphi)u = \varphi(au)$，则 F 作用在 $\mathrm{Hom}_F(U,V)$ 上. 若令 $\varphi \mapsto \varphi(1)$ 是 $\mathrm{Hom}_R(R,M)$ 到 M 的映射，则 $\mathrm{Hom}_R(R,M) \cong M$.

类似于群的合成列，我们说 R 模 M 的一个子模列 $M = M_0 \geqslant M_1 \geqslant \cdots \geqslant M_r = 0$ 是一个合成列，如果这个列中任意相邻两模的商模 $M_i/M_{i+1}\{i = 0, 1, \cdots, r-1\}$ 是单模，那么 M_i/M_{i+1} 称为 M 的合成因子. 一个模不一定有合成列. 例如，Z 模没有合成列，因为不存在 Z 的一个子模列使其最后一项终止于 0.

对于一个有合成列的模来说，Jordan-Holder 定理仍然成立，且模 M 的子模和商模的合成因子也是 M 的一个合成因子.

7.2 主理想整环上的自由模

7.2.1 自由模

设 R 是一个环和 R^n 是 n 元序组 $(x_1,x_2,\cdots,x_n)$ 的集合，$x_i \in R$ 作为 n 维向量空间 R^n 的我们熟悉的结构，我们将如下面那样引进加法，R^n 的 0 元以及用 R 的元来乘的乘法.

$$(x_1,\cdots,x_n) + (y_1\cdots,y_n) = (x_1+y_1,\cdots,x_n+y_n),$$

$$a(x_1,\cdots,x_n) = (ax_1,\cdots,ax_n),$$

显然 $(R^n,+,0)$ 是一个 Abel 群，且对 R^n，模的条件(1)～(4)都成立，因此 R^n 是环 R 上的模. 特别地，当 $n=1$ 时，R^1 就是通常作为左 R 模的 R. 令 $e_i = (0,\cdots,0,1,0,\cdots,0)$，那么 $x_ie_i = (0,\cdots,0,x_i,0,\cdots,0)$. 因此这 n 个元 e_i 生成了 R^n，它是一个 R 模.

有限维向量空间一定有基，并且其完全由它的维数所决定. 下面介绍与此类似的自由模概念.

定义 7.2.1　设 R 是环，M 是 R 模. M 的一个子集 S 称为 R **线性无关**的，如果对 S 的任意有限子集$(x_1,\cdots,x_n)$ 以及 $r_i \in R, 1 \leqslant i \leqslant n$，只要 R 线性组合 $r_1x_1+\cdots+r_nx_n=0$，就有 $r_1=\cdots=r_n=0$，否则称为 R **线性相关**的.

定义 7.2.2　R 线性无关的生成元集 S 称为 M 的一个基. 有基 S 的模称为 S 上的**自由模**，简称自由模. 其中元素的基数称为**自由模的秩**，记为 $r(M)$.

由自由模的定义可知，R 模 M 是 S 上的自由模当且仅当 M 中每个元可唯一写为 S 中元的 R 线性组合. 对有限生成自由模，我们还可给出更多一些的描述：

定理 7.2.1　设 R 是环，对 R 模 M 来说，下列条件等价：

1) M 是具有基 $\{x_1,\cdots,x_n)$ 的自由模；

2) $M=Rx_1\oplus\cdots\oplus Rx_n$；

3) $M\cong R\oplus\cdots\oplus R$.

证明　1)$\Rightarrow$2)　设 X 是 M 的一个基，由 X 的线性无关性知，M 中元素可唯一写为 $r_1x_1+\cdots+r_nx_n$，其中 $r_i\in R, 1\leqslant i\leqslant n$，故 M 是 $Rx_1,\cdots,Rx_n$ 的直和.

2)$\Rightarrow$3)　设 X 是 M 的一个基，$x\in X$，则易验证 $f: r\mapsto rx$ 是 R 到 Rx 的一个 R 模满同态. 由 X 的线性无关性知，若 $rx=0$，则 $r=0$，故 f 是单射. 因而 R, Rx 作为 R 模，我们有 $R\cong Rx$. 由此 $R\oplus R\cong Rx\oplus Rx$. 归纳可得 $M\cong R\oplus\cdots\oplus R$.

3)$\Rightarrow$1)　设 $M\cong R\oplus\cdots\oplus R$，令 $r_i=\{0,\cdots,0,1,0,\cdots,0\}$，其中第 i 个分量是 R 的单位元 1，其余分量为 0. 于是$\{r_i \mid i=1,\cdots,n\}$ 是 $R\oplus\cdots\oplus R$ 的一个基，$\{r_i \mid i=1,\cdots,n\}$ 在同构映射下的像即为 M 的一个基，故 M 是秩为 n 的自由模. 此结论可推广到任意的自由模.

7.2.2　自由模和矩阵

在线性代数中，对于给定的有限维向量空间，它的任意一个基中所含向量的个数都相等. 这个事实对一般环上的自由模并不一定成立. 但是我们有下面的结果：

定理 7.2.2　设 R 是环，I 是 R 的理想且 R/I 是除环，M 是自由 R 模，则 M 的任何两个基有相同的基数.

证明　令 $K=IM=\{rx \mid r\in I, x\in M\}$，则 K 是 M 的子模. 对于 $x+K\in M/K$，$x+I\in R/I$，令$(r+I)(x+K)=rx+K$，则 M/K 是 R/I 上的向量空间. 设 B 是 M 的基，令 $\overline{B}=\{\overline{b}\mid b\in b+K\}$，则对于每个 $\overline{x}\in M/K$，我们有 $\overline{x}=\sum\limits_{i=1}^{k}r_ib_i+k=\sum\limits_{i=1}^{k}(r_i+I)\overline{b_i}$，$\forall\, r_i\in R, b_i\in B, \overline{b}\in\overline{B}$，故 $\overline{B}$ 生成 M/K. 如果 $\sum\limits_{i=1}^{k}(r_i+I)\overline{b_i}=\sum\limits_{i=1}^{k}r_ib_i+K=\overline{0}=K$，则 $\sum\limits_{i=1}^{k}r_ib_i\in K$. 于是存在 $s_1,\cdots,s_k\in I$，使得 $\sum\limits_{i=1}^{k}r_ib_i=\sum\limits_{i=1}^{k}s_ib_i$，也即 $\sum\limits_{i=1}^{k}(r_i-s_i)=0$.

由此对任意 $i, s_i = r_i$,因而 $\overline{r_i} = r_i + I = \overline{0}$. 这说明 $\overline{B}$ 是 M/K 在 R/I 上的基.

特别地,$b \mapsto \overline{b}$ 是 B 到 $\overline{B}$ 的一一对应,从而 $\dim_{R/I}(M/K) = |B|$.

推论 7.2.3 设 R 是交换环,M 是自由 R 模,则 M 的任何两个基有相同的基数.

推论 7.2.4 设 R 是交换环,M 为有限生成的自由 R 模,则 M 的任意基所含的元素个数都相等.

一个自由模的子模不一定是自由模.

例 7.2.1 令 $R = \mathbf{Z}_4$,则 $M = 2R = \{[0],[2]\}$ 是 R 的理想,而且 M 是自由 R 模 R 的子模;然而它不是一个自由 R 模,这是因为 $|M| = 2$, $|R| = 4$,由此得 M 没有基.

定理 7.2.5 设 R 是主理想环,M 是自由 R 模,N 是 M 的子模,则 N 也是自由模,且 $r(N) \leqslant r(M)$.

定理 7.2.7 如果 R 是交换环,$R^m \cong R^n$,那么 $m = n$.

证明 令 $\{e_i \mid 1 \leqslant i \leqslant m\}$,$\{f_j \mid 1 \leqslant j \leqslant n\}$ 是 M 的两个基,a_{ji} 和 $b_{ij} \in R$,那么

$$f_j = \sum_1^n a_{ji} e_i, e_i = \sum_1^m b_{ij} f_j,$$

代入得 $f_j = \sum_{i=1, j'=1}^{n,m} a_{ji} b_{ij'} f_{j'}, e_i = \sum_{j=1, i'=1}^{m} b_{ij} f_j a_{ji'} e_{i'}$,

因为 f 和 e 组成基,我们有

$$\sum_{i=1}^n a_{ji} b_{ij'} = \begin{cases} 1, \text{当 } j = j' \text{ 时}, \\ 0, \text{当 } j \neq j' \text{ 时}, \end{cases} \sum_{j=1}^m b_{ij} a_{ji'} = \begin{cases} 1, \text{当 } i = i' \text{ 时}, \\ 0, \text{当 } i \neq i' \text{ 时}, \end{cases} \tag{7.1}$$

这里 $j, j' = 1, 2, \cdots, m, i, i' = 1, 2, \cdots, n$. 现在假设 $m < n$,再考虑两个 $n \times n$ 矩阵:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1} & a_{m1} & \cdots & a_{mn} \\ 0 & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 0 \end{pmatrix} \qquad B = \begin{pmatrix} b_{11} & \cdots & b_{1m} & 0 & \cdots & 0 \\ b_{21} & \cdots & b_{2m} & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{n1} & \cdots & b_{nm} & 0 & \cdots & 0 \end{pmatrix}$$

(7.1)式相当于矩阵条件 $BA = I$,因为 r 是可换的,这可导出 $AB = I$. 然而从矩阵 A 和 B 的形式来看 AB 的最后 $n - m$ 行是 0,所以 $AB \neq I$,这个矛盾说明 $m \geqslant n$,从对称性可得 $n \geqslant m$,所以 $m = n$.

如果$(e_1, \cdots, e_n)$和$(f_1, \cdots, f_n)$都是基,而 $f_j = \sum_{i=1}^n a_{ji} e_i, e_i = \sum_{j=1}^n b_{ij} f_j$,那么

$AB = I$,因此 A 与 B 可逆,即 $A,B \in Ln(R)$ 表示元在 R 中的 $n \times n$ 可逆矩阵群. 反之,假设$(e_1, \cdots, e_n)$是一个基,而 $A \in Ln(R)$,定义

$$f_j = \sum_{i=1}^{n} a_{ji} e_i, 1 \leqslant j \leqslant n,$$

那么$(f_1, \cdots, f_n)$也是一个基,因为 $BA = I$,就得到对一切 $h, dh = 0$,于是$(f_1, \cdots, f_n)$是一个基. 这个结果告诉我们,如果有可换环 R 上的自由模的一个有序基$(e_1, \cdots, e_n)$,那么我们就获得所有的序基$(f_1, \cdots, f_n)$. 这是应用矩阵 $A \in Ln(R)$ 于(e_i)上而得到的,即取 $f_j = \sum\limits_{i=1}^{n} a_{ji} e_i, A = (a_{ij})$.

我们现在不限制 R 为可换的,对任何 m,n 考虑从 R^m 到 R^n 的(模)同态的加法群 $\mathrm{Hom}_R(R^m, R^n)$,我们分别选择 R^m 和 R^n 的基$(f_1, \cdots, f_n)$,如果 $\eta \in \mathrm{Hom}_R(R^m, R^n)$,我们列出

$$\eta(e_1) = a_{11} f_1 + a_{12} f_2 + \cdots + a_{1n} f_n$$

$$\eta(e_2) = a_{21} f_1 + a_{22} f_2 + \cdots + a_{2n} f_n$$

$$\cdots\cdots\cdots$$

$$\eta(e_m) = a_{m1} f_1 + a_{m2} f_2 + \cdots + a_{mn} f_n$$

称这个 $m \times n$ 矩阵 $A = (a_{ij})_{mn}$ 为关于基$(e_1, \cdots, e_n)$,$(f_1, \cdots, f_n)$的矩阵. 同态 η 由它相对于基$\{e_i\}$,$\{f_i\}$所确定,$x = (x_1, \cdots, x_m) = \sum x_i e_i$,那么

$$\eta(x) = \eta(\sum x_i e_i) = \sum x_i \eta(e_i) = \sum_{i,j} x_i a_{ij} f_j$$

是映射

$$(x_1, \cdots, x_m) \to (y_1, \cdots, y_n),$$

这里

$$y_i = \sum_{j=1}^{m} x_i a_{ij}, j = 1, \cdots, n.$$

一般地,如果 $A=(a_{ij})$和 $B=(b_{ij})$都是矩阵,我们定义和 $A+B=(a_{ij}+b_{ij})$,如果 $A=(a_{ij})$是一个 $m\times n$ 矩阵和 $B=(b_{ij})$是一个 $n\times q$ 矩阵,那么我们定义积 $P=AB$ 是 $m\times q$矩阵,它的$(i, ,k)$元$(1\leqslant i\leqslant m, 1\leqslant k\leqslant q)$是由公式 $p_{ik} = \sum\limits_{j=1}^{n} a_{ij} b_{jk}$ 给出的.

以 R 维元的 $m\times n$ 矩阵构成的集合 $M_{m\times n}(R)$是一个群,

推论 7.2.8　在映射 $\eta \to A$ 之下,这个群同构于 $\mathrm{Hom}(R^{(m)}, R^{(n)})$,这里 A 是关于 $R^{(m)}$ 的基$\{e_i\}$和 $R^{(n)}$ 的基$\{f_i\}$的矩阵.

7.2.3 模的直和

模的直和概念如同群的直和概念一样,在模论研究中是相当重要的.

定义 7.2.3 设 R 是环, $n\in N, M_1,\cdots,M_n$ 是 R 模,作集合的积 $M_1\times M_2\times\cdots\times M_n$. 在这一集合中定义二元运算

$$(m_i,\cdots,m_n)+(m_1',\cdots,m_n')=(m_1+m_1',\cdots,m_n+m_n').$$

易验证积是一个加群,其零元为$(0,\cdots,0)$,$(m_1,\cdots,m_n)$ 的负元为$(-m_1,\cdots,-m_n)$.

定义 7.2.4 对于 $\forall r\in R$,定义 $r(m_1,\cdots,m_n)=(rm_1,\cdots,rm_n)$,则 $M_1\times\cdots\times M_n$ 关于以上定义的运算是一个 R 模,称具有这个运算的加氏积 $M_1\times\cdots\times M_n$ 为 $M_1,\cdots,M_n$ 的(外)直和,表示为: $M_1\oplus\cdots\oplus M_n$. 显然

$$M_1\oplus\cdots\oplus M_n\cong M_{p(1)}\oplus\cdots\oplus M_{p(n)},\forall p\in S_n,$$

故直积与因子的顺序无关.

$M=M_1\oplus\cdots\oplus M_n$ 具有下列性质:

1) 对 $\forall m\in M$ 有子模 N_i 同构于 M_i,特别地,$K_i=\{(0,\cdots,m_i,\cdots,0)\mid m_i\in M_i\}$,这里 m_i 在第 i 个位置上,而且 M/K_i 同构于其余 M_j 的直和.

2) 对 $\forall m\in M$,m 有唯一的表达式 $m=k_1+\cdots+k_n$,其中 $k_i\in K_i$. 若 $m=(m_1,\cdots,m_n)$,则 $k_i=(0,\cdots,m_i,\cdots,0)$ (其中 m_i 出现在第 i 个位置上). 现在设 $M_1,\cdots,M_n$ 是模 M 的子模,故 $M_1+\cdots+M_n=\{m_1+\cdots+m_n\mid m_i\in M_i,1\leqslant i\leqslant n\}$.

3) 假如给定了由 M_i 到模 N 的同态 $\eta_i,1\leqslant i\leqslant n$,那么我们用

$$\eta:(x_1,x_2,\cdots,x_n)\to\sum_1^n\eta_i(x_i),$$

来定义从 $\oplus M_i$ 到 N 的映射 η,因为

$$\begin{aligned}\eta(x_1y_1+\cdots+x_ny_n)&=\sum_1^n\eta_i(x_i+y_i)=\sum_1^n\eta_i(x_i)+\sum_1^n\eta_i(y_i)\\&=\eta(x_1,x_2,\cdots,x_n)+\eta(y_1,y_2,\cdots,y_n),\end{aligned}$$

$$\eta(ax_1,ax_2,\cdots,ax_n)=\sum_1^n\eta(ax_i)=\sum_1^n a(\eta x_i)=a\sum_1^n\eta_i(x_i),$$

所以 η 是$\oplus M_i$ 到 N 的一个同态.

定理 7.2.9 设 $M_1,\cdots,M_n$ 是模 M 的子模,则下列条件是等价的:

1) $M_1\oplus\cdots\oplus M_n\cong M_1+\cdots+M_n$,同构映射为 $f:(m_1,\cdots,m_n)\mapsto m_1+\cdots+m_n$;

2) 对于 $\forall j\in\{1,2,\cdots,n\},M_j\cap(M_1+\cdots+M_{j-1}+M_{j+1}+\cdots+M_n)=0$;

3) $M_1+\cdots+M_n$ 中每个元可唯一表示为 $m_1+\cdots+m_n$ 的形式,其中 $m_i\in M_i$.

证明 1)⇒2) 若对某个 j, 2)不成立. 令

$$0 \neq m_j \in M_j \cap (M_1 + \cdots + M_{j-1} + M_{j+1} + \cdots + M_n),$$

则
$$m_j = m_1 + \cdots + m_{j-1} + m_{j+1} + \cdots + m_n,$$
于是
$$f(m_1 + \cdots + m_{j-1} + m_{j+1} + \cdots + m_n) = 0,$$
从而 $\mathrm{Ker} f \neq 0$,矛盾.

2)⇒3) 对 j 用归纳法. 当 $j=1$ 时,结论显然成立;设对 $j-1$ 结论成立,对于 $m \in M_1 + \cdots + M_j$,若 $m = m_1 + \cdots + m_j = m_1' + \cdots + m_j'$,$m_i$, $m_i' \in M_i$,$1 \leqslant i \leqslant j$,则我们有 $m_j'm_j = (m_1 + \cdots + m_{j-1}) - (m_1 + \cdots + m_{j-1}) \in M_j \cap (M_1 + \cdots + M_{j-1})$. 由假设知,$m'_j - m_j = 0$,即 $m'_j = m_j$,由此得 $m_1 + \cdots + m_{j-1} = m_1' + \cdots + m'_{j-1}$,由归纳假设得,$m'_i = m_i$,其中 $1 \leqslant i \leqslant j-1$. 于是对所有 i 来说,$m'_i = m_i$,结论对 j 成立.

3)⇒1) 注意到 f 是模的满同态,又由 3) 易知,f 是单射,故 f 是同构映射.

如果 M_i 是满足条件 1) 和 2) 的 M 的子模,这个定理容许把模 M 与 $M_1 \oplus \cdots \oplus M_n$ 认为同一的. 这时说 M 是它的子模 M_i 的(内) 直和,并写成 $M = M_1 \oplus \cdots \oplus M_n$,只要子模 M_i 适合条件 1) 与 2).

如果说 M_i 的集合,$1 \leqslant i \leqslant n$,满足条件 2),那么我们说 M 的这些子模是无关的,这时当且仅当形如 $\sum_1^n x_i = 0$ 的关系成立时,$x_i \in M_i$,$x_i = 0$.

这个无关条件比条件 $M_i \cap M_j = 0$,$i \neq j$ 强,甚至比 $M_i \cap (\bigcup_{j \neq i} M_j) = 0$ 还要强.

定义 7.2.5 设 $M_1, \cdots, M_n$ 是模 M 的子模,若 R 模 $M = M_1, \cdots, M_n$ 满足定理7.2.9 中的 2) 和 3),则称 M 是子模 $M_1, \cdots, M_n$ 的**内直和**.

因为子模 $M_1, \cdots, M_n$ 的内直和与它们的外直和是同构的,所以我们在使用时常常将它们视为相同的,并将 $M_1, \cdots, M_n$ 的内直和也记为 $M_1 \oplus \cdots \oplus M_n$. 有时我们用 nM 表示 n 个模 M 的直和,有时也用 M^n 表示 n 个模 M 的直和.

定义 7.2.6 一个模称为**半单**的,若它是单模的直和.

模 M 的子模 N 称为 M 的一个**直和因子**,若存在 M 的一个子模 N',使得 $M = N \oplus N'$.

若模 M 的任何一个子模都是 M 的一个直和因子,则称 M 是**完全可约模**.

易见半单模与完全可约模是一致的.

7.3 主理想整环上的有限生成模

与讨论其他代数系统问题类似,给定一个环 R,决定所有可能的 R 模是模论研究的一个基本问题. 对于一般的环,这个问题过于困难,而对一些特殊的环我们可给出这个

问题的完全解答.

当 R 是域时,域上的模就是线性空间,其结构完全由它的维数所决定.

当 R 是有单位元的交换环时,R 上的有限生成自由模完全由它的秩所决定.

我们准备讨论当 R 是主理想环时,R 上的有限生成模的结构问题. 这个结论的两个直接应用: 研究主理想整环上有限生成模和这个理论在有限 Abel 群与线性变换上的应用,以及复数域上任何 n 阶方阵与 Jordan 标准形相似的理论证明.

7.3.1 主理想环上有限生成自由模的性质

定理 7.3.1 D 是一个主理想环,D^n 是 D 上的秩为 n 的自由模,那么 D^n 的任意子模 K 是具有 $m \leqslant n$ 个元的基的自由模.

证明 首先约定仅有 0 的模是"秩为 0 的自由模",当然当 $n=0$ 时,命题成立.

现在假设 $n=1$,那么 $D^{(1)}$ 与 D 一样,K 的任一子模是一个理想,因此 $K=(f)$,如果 $f=0$,我们有 $K=0$,否则 $f \neq 0$,因为 D 是整环,$af=0(a \in D)$,可以得 $K=0$. 于是 K 是以 f 为基的自由模,这证明了当 $n=1$ 时命题成立.

假定 $n>1$,对 n 进行归纳法,假定$\{e_1,\cdots,e_n\}$ 是 $D^{(n)}$ 的一个基,$D^{(n-1)}$ 是由$\{e_2,\cdots,e_n\}$ 所生成的子模,那么 $D^{(n-1)}$ 是秩为 $n-1$ 的自由模,$D^{(n)}/D^{(n-1)}$ 是具有基 $\overline{e}_1=e_1+D^{(n-1)}$ 的一个子模. $\overline{K}=(k+D^{(n-1)})/D^{(n-1)}$ 是 $\overline{D}^{(n)}=D^n/D^{(n-1)}$ 的一个子模. 如果$\overline{K}=0$,$k+D^{(n-1)}=D^{(n-1)}$,所以 $k \in D^{(n-1)}$,那么由归纳假定,定理得证.

假定$\overline{K} \neq 0$. 对 $n=1$,$D^{(1)}$ 有一个基 $\overline{f}_1=f_1+D^{(n-1)}$, 又因为$\overline{K}=(k+D^{(n-1)})/D^{(n-1)}$. 可以选 $f_1 \in K$. 归纳假设应用于 $D^{(n-1)}$ 的子模 $K \cap D^{(n-1)}$ 可得,如果 $K \cap D^{(n-1)} \neq 0$,那么这个子模有一个基$\{f_1,\cdots,f_m\}$,$0<m-1 \leqslant n-1$,那么$\{f_1,\cdots,f_m\}$ 是 K 的一个基. 令 $y \in K$,就有 $\overline{y}=y+D^{(n-1)} \in K$. 而 $\overline{y}=b_1\overline{f}_1$,$b_1 \in D$,这就是 $y-b_1f_1 \in D^{(n-1)}$,因为我们选取 $y \in K$,$y-b_1f_1 \in K$,所以 $y-b_1f_1 \in K \cap D^{(n-1)}$,从而有 $y-b_1f_1=b_2f_2+\cdots+b_mf_m$. 因此 $y=\sum\limits_1^m b_jf_j$,这些 f_j 生成 K.

再假设 $\sum\limits_1^m b_jf_j=0$,那么 $b_1\overline{f}_1=\sum\limits_2^m b_k\overline{f}_k=0$,因为 $\overline{f}_i$ 是 $\overline{D}^{(n)}$ 的一个基,所以 $b_1=0$,并且因 f_k,$k \geqslant 2$ 构成 $K \cap D^{(n-1)}$ 的一个基,$\sum\limits_2^m b_kf_k=0$, 即 $b_k=0$,因此这些 f 构成 K 的一个基. 如果 $K \cap D^{(n-1)}=0$,同样可以证明 f_j 是 K 的一个基.

因为任何域是一个主理想整环,所以上面的定理可以应用于 D 是一个域 F 上的 n 维向量空间,V 的任一子空间是有限维的,它的维数 $m \leqslant n$.

现在回到 $M \cong D^{(n)}/K$. 可以断定 K 有一个 $m \leqslant n$ 元的基,假定$\{f_1,\cdots,f_m\}$是子模 K 的生成元,这里 m 可以大于 n,现在用基$\{e_1,\cdots,e_n\}$表示这些生成元.

$$f_1 = a_{11}e_1 + a_{12}e_2 + \cdots + a_{1n}e_n$$
$$f_2 = a_{21}e_1 + a_{22}e_2 + \cdots + a_{2n}e_n$$
$$\cdots \quad \cdots \quad \cdots \quad \cdots$$
$$f_m = a_{m1}e_1 + a_{m2}e_2 + \cdots + a_{mn}e_n$$

这些关系的 $m\times n$ 矩阵 $A=(ak_i)$ 称为有序生成元$\{f_1,\cdots,f_m\}$用有序基$\{e_1,\cdots,e_n\}$来表示的关系矩阵.

定理 7.3.2 D 是主理想环,M 是有限生成模,N 是 M 的子模,则 N 也是有限生成模.

7.3.2 Noetherian 环

定义 7.3.1 左 R 模 M 称为**满足升链条件**,若 $M_1\subseteq M_2\subseteq\cdots$ 是 M 的一个子模升链,则存在正整数 n 使得对所有的 $k\geqslant n$ 都有 $M_k=M_n$.

环 R 称为 **Noetherian 环**,若 R 作为 R 模满足升链条件.

定理 7.3.3 设 R 是环,M 是左 R 模,则下列条件是等价的:

1) M 满足升链条件;

2) 每个由 M 的子模构成的非空集合在包含关系下有一个极大元;

3) M 的每个子模是有限生成的,特别地,M 是有限生成的.

证明 1)⇒2) 设 M 满足升链条件,T 是任意一个由 M 的子模构成的非空集合. 任取 $M_1\in T$,若 M_1 是极大元,则 2) 成立. 若 M_1 不是极大元,则存在 $M_2\in T$ 使得 $M_1\subset M_2$. 若 M_2 是极大元,则 2) 成立,否则这个过程可继续下去,若 2) 不成立,则我们就得到 M 的一个严格升链,它含有无限多个不同子模,这与 1) 矛盾,因而 2) 成立.

2)⇒3) 设 2) 成立. 令 N 是 M 的子模,T 是 N 的所有有限生成子模构成的集合. 因为$\{0\}\in T$,故 T 非空. 由 2) 知 T 含有一个极大元 N. 若 $N'\neq N$,令 $x\in N-N'$,因为 $N'\in T$,故 N' 是有限生成的,因而由 N' 和 x 生成的子模也是有限生成的. 这与 N' 的极大性相矛盾,故 N 是有限生成的.

3)⇒1) 设 3) 成立,$M_1\subseteq M_2\subseteq\cdots$ 是 M 的一个子模升链. 令 $N=\bigcup_{i=1}^{\infty}M_i$,由 3) 可知 N 是有限生成的. 设其一个生成元集为$\{e_1,e_2,\cdots,e_r\}$,于是对每个 i 来说,e_i 必属于某个 M_j,令 $y=\max\{j_1,j_2,\cdots,j_r\}$,则对任意 i 来说,$e_i\in M_y$,由此得 $N\subseteq M_y$,这必须有 $M_y=N=M_u$,任意 $u\geqslant y$,则 M 满足升链条件,故 1) 成立.

推论 7.3.4 设 D 是主理想环,则每个由 D 的理想构成的非空集合在包含关系下存在一个极大理想,且 D 是 Noetherian 环.

7.3.3 主理想整环上的矩阵的等价

主理想整环 D 上的两个 $m\times n$ 矩阵 A, B,如果在 $M_m(D)$上存在一个可逆矩阵 P,在 $M_m(D)$上存在一个可逆矩阵 Q 使 $B=PAQ$,那么说 A, B 是等价的,显然这定义了

元在 D 上的 $m\times n$ 矩阵集合 $M_m(D)$ 里的一个等价关系，我们现在考虑从与已知矩阵等价的矩阵中挑选一个特别简单的“正规”型的问题.

综上给出下面的结论：

定理 7.3.5 如果 $A\in M_m(D)$，D 是一个理想整环，当 $i\leqslant j$ 且 $d_i\neq 0$ 时 $d_i\mid d_j$，那么 A 等价于对角形矩阵 $\mathrm{diag}(d_1, d_2, \cdots, d_r, 0, \cdots, 0)$.

将得到把 A 变成对角形矩阵的矩阵 P 与 Q，它们是即将定义的特殊矩阵的积，不指定 $m\times n$ 或 $n\times n$，首先引进某些元在 D 上的可逆阵，称为初等矩阵.

首先设 $P_{ij}=1-e_{ii}-e_{jj}+e_{ij}+e_{ji}$. 因为 $P_{ij}^2=1$，所以这个矩阵是可逆的，且

$$P_{ij}^{-1}=P_{ij}.$$

其次设 u 是 D 的一个可逆元，令 $D_i(u)=1+(u-1)e_{ii}$，即 $D_i(u)$ 是一个对角矩阵，它的对角线上第 1 个是 u，其他都是 1. 于是 $D_i(u)$ 可逆，且

$$D_i^{-1}(u)=D_i(u^{-1}).$$

最后设 $b\in D$，$i\neq j$，令 $T_{ij}(b)=1+be_{ij}$，在 (i, j) 位置上是 e_{ij}，其他位置上都是 0 的矩阵. 因为 $T_{ij}(b)\,T_{ij}(-b)=(1+be_{ij})(1-be_{ij})=I$，所以 $T_{ij}(b)$ 可逆，且

$$T_{ij}^{-1}(b)=T_{ij}(-b).$$

我们称矩阵 P_{ij}、$D_i(u)$、$T_{ij}(b)$ 分别为 Ⅰ，Ⅱ，Ⅲ 型初等矩阵.

定义 7.3.2 具有如定理 7.3.5 中给出的等价于 A 的对角矩阵叫做 A 的一个**标准型**，一个标准型的对角线上的元叫做 A 的**不变因子**.

定理 7.3.6 设 $A\in M_{mn}(D)$，A 的秩为 r，对每一 $i\leqslant r$，令 Δ_i 是所有 i 级子式的最大公因子，A 的任一不变因子组与 $d_1=\Delta_1, d_2=\Delta_2\Delta_1^{-1}, \cdots, d_r=\Delta_r\Delta_{r-1}^{-1}$ 至多相差一个单位因子. 显然 $\Delta_i\neq 0$，以及 $\Delta_{i-1}\mid\Delta_i$.

证明 令 $Q=(q_{ij})\in M_{mn}(D)$，于是 QA 的在 (k,i) 位置的元是 $\sum_j q_{kj}a_{ji}$，这表明 QA 的行是系数在 D 里的 A 的行的线性组合. 因此 QA 的 i 级子式是 A 的 i 级子式的线性组合，从而 A 的所有 i 级子式的最大公因子是 QA 的所有 i 级子式的最大公因子的一个因子.

同样，当 $P\in M_n(D)$，AP 的列是 A 的列的线性组合，A 的所有 i 级子式的最大公因子是 AP 的所有 i 级子式的最大公因子的一个因子，给定这两个事实和根据等价关系的对称性，可知假使 A 与 B 等价，A 和 B 的 i 级子式的最大公因子是相同的，现在令 $B=\mathrm{diag}(d_1, d_2, \cdots, d_i, 0, \cdots, 0)$ 是 A 的一个标准型，于是从可除性 $d_i\mid d_j$（当 $i<j$ 时）可以导出 B 的 i 级子式的一个最大公因子是 $\Delta_i=d_1d_2\cdots d_i$，结论成立.

推论 7.3.7 两个 $m\times n$ 矩阵是等价的当且仅当它们有相同的不变因子.

7.4 主理想整环上有限生成模的结构

7.4.1 扭模

定义 7.4.1 设 D 是主理想环，M 是 D 模，称 M 的一个元素 x 为**扭元素**，如果 D 中存在非零元素 d 使得 $dx=0$，否则称 x 为自由的. 若 M 中每个元素都是扭元素，则称 M 为**扭模**. 若 M 的每个非零元素是自由的，则称 M 为**无扭模**.

例 7.4.1 设 D 是主理想环，M 是 D 模，若 $x\in M$ 是一个扭元素，令

$$A(x)=\{r\in D\mid rx=0\},$$

则易验证 $A(x)$ 是 D 的理想. 因为 D 是主理想环，故存在 $rx\in D$ 使得 $A(x)=(rx)$.

定义 7.4.2 称 rx 为 D 的**阶**，记作 $|x|$，有时也把 rx 称为循环模 R_x 的阶. $A(x)=(rx)$ 称为 x 的**阶理想**(或**零化理想**).

需要注意的是，对于 x 的阶来说，x 的阶 rx 在和 D 相差一个单位的情况下是唯一确定的，我们知道，若 u 是 D 的一个单位，即 u 是 D 的一个可逆元，则 $(rxt)=(urx)$，也即 rx 的任何一个相伴元都是 x 的阶. 对于被 r 零化的子模来说，$M(r)$ 中每个元素的阶整除 r.

定理 7.4.1 设 M 是主理想环 D 上的有限生成模，则 M 是自由的当且仅当 M 是无扭的.

证明 *必要性* 不失一般性，设 $M\neq 0$，$x_1,\cdots,x_n$ 是 M 的一个基. 如果 M 不是无扭的，则存在 M 的非零元素 x 以及 D 的非零元 r 使得 $rx=0$，设 $x=r_1x_1+\cdots+r_nx_n$，$r_i\in D$，则 $0=rx=rr_1x_1+\cdots+rr_nx_n$，因为 $x_1,\cdots,x_n$ 是自由模 M 的基，故对所有的 i 来说，$rr_i=0$. 又 $r\neq 0$，D 是主理想环，从而对 $\forall i$，$r_i=0$，由此得 $x=0$，这与所设矛盾. 故 M 是无扭模.

充分性 设 $\{x_1,\cdots,x_n\}$ 是 M 的一个生成元集，$\{y_1,\cdots,y_m\}$ 是 $\{x_1,\cdots,x_n\}$ 的一个极大线性无关的子集. 令 N 是由这 m 个元生成的 M 的秩为 m 的自由子模，由 $\{y_1,\cdots,y_m\}$ 的极大性可知，对 $\forall x_i$ 来说，存在 $0\neq s_i,r_{ij}\in D$，使得 $s_ix_i+\sum\limits_j r_{ij}y_j=0$，故 $s_ix_i\in N$. 令 $s=\prod\limits_{i=1}^n s_i$，则对 $\forall X\in M$ 都有 $sX\in N$. 令 φ 是 M 到 N 的映射，其中 $\varphi(X)=sx$，验证可知 $\varphi\in \mathrm{Hom}_R(M,N)$. 又因为 m 是无扭的，故 $\mathrm{Ker}\varphi=0$. 于是 M 同构于 N 的一个子模，即得 M 的秩不超过 m 的自由模.

下面我们讨论主理想环上有限生成模. 主要讨论有限生成扭模，归结为：在有限生成模中，只有一个元生成的模是循环模. 我们将通过一系列引理来证明主理想环上有限生成模可分解为一个自由模与若干个循环子模的直和.

引理 7.4.2 设 R 是整环，T 是由 R 模 M 的所有扭元素构成的集合，则 T 是 M 的子模且 M/T 是无扭模.

证明 设 $x,y\in T$，$|x|=a$，$|y|=b$，则 $ab(x-y)=b$，$A(x)=(rx)$，$x-a\cdot by=0$. 又对于 $r\in R$，$a(rx)=r(az)=0$，故 T 是 M 的子模. 对于 $r\in R$，$x+T\in M/T$，若 $r(x+T)=rx+T=T$，则 $rx\in T$，于是 $A(rx)\neq 0$. 设 $0\neq s\in A(rx)$，则 $sr\cdot x=0$，由此得 $0\neq sr\in A(x)$，从而 $x\in T$，这说明 M/T 是无扭的. 子模 T 称为 M 的**扭子模**.

定理 7.4.3 设 D 是主理想环，M 是 R 模，F 是自由 R 模，$\varphi\in \mathrm{Hom}_R(M,F)$ 为满射，则 M 有一个与 F 同构的自由子模 E 使得 $M=E\oplus \mathrm{Ker}\varphi$.

证明 设 B 是 F 的一个基，对每个 $b_i\in B$，选择 $x_{b_i}\in M$ 使得 $\varphi(x_{b_i})=b_i$. 令 E 是由这些 x_{b_i}，$b_i\in B$，生成的 M 的子模. 如果 $\sum\limits_{i=1}^{k} r_i x_{b_i}=0$，$r_i\in D$，则

$$0=\varphi(\sum_{i=1}^{k} r_i x_{b_i})=\sum_{i=1}^{k} r_i x_{b_i},$$

于是，所有的 $r_i=0$. 这就是说，$\{x_{b_i}\in B\}$ 是线性无关的，从而 $\{x_{b_i}\in B\}$ 是 B 的一个基，即 B 是自由模且同构于 F. 对于 $x\in M$，若 $\varphi(x)=\sum_i r_i b_i$，$r_i\in D$，$b_i\in B$，则 $x-\sum_i r_i b_i\in \mathrm{Ker}\varphi$，于是 $M=E+\mathrm{Ker}\varphi$，又 $E\cap \mathrm{Ker}\varphi=0$，从而 $M=E\oplus \mathrm{Ker}\varphi$.

定理 7.4.4 设 D 是主理想环，M 是有限生成 R 模，T 是 M 的扭子模，则 M 有一个有限秩的自由子模 F 使得 $M=T\oplus F$，其中 F 的秩由 M 唯一确定.

证明 由引理 7.4.2 知 M/T 是无扭模，自然同态 η: $M\to M/T$ 是满射. 由定理 7.4.1 及引理 7.4.2 可知，存在 M 的一个与 M/T 同构的自由子模 F 使得 $M=T\oplus F$，若另有 M 的自由子模 F' 使得 $M=T\oplus F'$，则易知 $F'\cong M/T$，因而 $r(F)=r(F')$.

7.4.2 有限生成扭模

引理 7.4.5 设 M 是主理想整环 D 上的模，x 和 y 是 M 的扭元素，它们的阶分别是 r 和 s，且 $(r,s)=1$，则 $|x+y|=rs$.

证明 设 $|x+y|=n$，因为 $rs(x+y)=s\,rx+r\,sy=0$，故 $rs\in A(x+y)=(n)$，从而 $n\mid rs$，又 $rn(x+y)=(rn)y=0$，$sn(x+y)=snx=0$，于是有 $s\mid rn$，$r\mid sn$，又由于 $(r,s)=1$，故 $s\mid n$，$r\mid n$，从而 $rs\mid n$，于是 $n=prs$，其中 p 为 D 的单位，这时 $(n)=(rs)$，即 $|x+y|=rs$.

引理 7.4.6 设 D 是主理想环，M 是指数为 r 的扭模，则 M 必有阶为 r 的元素.

证明 主理想环是唯一分解环，则令 $d=p_1^{e_1}\cdots p_k^{e_k}$ 是 d 在 D 中的标准分解式，其中 p_i 是素元，$e_i>0$ 是整数. 再令 $d_i=d/p_i$，则 d 不整除 d_i. 于是对 $\forall i$，存在 $x_i\in M$，使得 $d_ix_i\neq 0$，又令 $y_i=(d/p_i^{e_i})x_i$，则有 $p_i^{e_i}y_i=0$，但是 $p_i^{e_i-1}y_i=d_ix_i\neq 0$，因而对任意的 i，

$|y_i| = p_i^{e_i}$. 最后令 $x = y_1 + \cdots + y_k$,由上引理得 $|x| = r$.

引理 7.4.7 设 D 是主理想环,$M = Dy$ 是阶为 r 的循环模,$s \in D$,则有

1) $M(s) = D((r/(r,s))y) \cong D/(r,s)$; 2) $sM = D((r,s)y) \cong D/(r(r,s))$.

也就是说,$M(s)$ 是阶为 (r,s) 的循环模,sM 是阶为 $r/(r,s)$ 的循环模.

证明 由定义有 $M(s) = \{uy \mid u \in D, suy = 0\}$. 又 $suy = 0 \Leftrightarrow r \mid su \Leftrightarrow r/(r,s) \mid s/(r,s)u \Leftrightarrow r/(r,s) \mid u \Leftrightarrow u$ 是 $r/(r,s)$ 的倍数,于是 $M(s) = D((r/(r,s))y)$. 令 φ 是 D 到 $M(s)$ 的映射,其中 $\varphi(x) = (xr/(r,s))y, x \in D$,易验证 φ 是满同态,且 $\mathrm{Ker}\varphi = (r,s)$,由同态基本定理即得.

类似可证 2).

引理 7.4.8 设 $W = Rz$ 是 r_0 阶循环模,M 是指数整除 r_0 的有限生成 R 扭模,N 是 M 的子模,$\varphi \in \mathrm{Hom}_R(N,W)$,则由 φ 可得到模同态 η,使 $\eta \in \mathrm{Hom}_R(M,W)$.

证明 设 $\{x_1, \cdots, x_k\}$ 是 M 的一个生成元集,令 $N_1 = N + Rx_1, N_2 = N_1 + Rx_2, \cdots, N_k = N_{k-1} + Rx_k = M$. 由归纳法我们只需证由 φ 可得到 N_1 到 W 的一个模同态 φ_1 即可.

设 $x_1 + N \in N_1/N$ 的阶为 s,则由题设可知 $s \mid r_0$,不妨设 $r_0 = st$,由于 $sx_1 \in N$,从而 $\varphi(sx_1) \in W$,且 $t\varphi(sx_1) = \varphi(tsx_1) = \varphi(r_0x_1) = 0$. 这说明 $|\varphi(sx_1)| \mid t$;

另一方面,存在 $u \in R$ 使得 $\varphi(sx_1) = uz$,故 $tuz = 0$. 由此得 $r_0 \mid tu$,又 $r_0 = ts$,故 $s \mid u$. 我们不妨设 $u = sv$,从而有 $\varphi(sx_1) = s(uz)$. 令 $z_0 = vz \in W$,定义 $N \oplus R$ 到 W 的映射

$$\varphi' : (y,r) \mapsto \varphi(y) + rz_0,$$

其中 $y \in N, r \in R$,可以验证 $\varphi' \in \mathrm{Hom}_R(N \oplus R, W)$.

定义 $N \oplus R$ 到 N_1 的映射 $\varphi'' : (y,r) \mapsto y + rx_1$,则 $\varphi'' \in \mathrm{Hom}_R(N \oplus R, N_1)$.

我们来证明 $\mathrm{Ker}\varphi'' \subseteq \mathrm{Ker}\varphi'$. 设 $(y,r) \in \mathrm{Ker}\varphi''$,则 $rx_1 = -y \in N$,于是 $r(x_1 + N) = N$,因而 $s \mid r$. 不妨设 $r = -sd$,则 $y = sdx_1, (y,r) = (sdx_1, -sd)$,于是我们有

$$\varphi'(y,r) = \varphi'(sdx_1, -sd) = \varphi(sdx_1) - sdz_0 = \varphi(sdx_1) - d\varphi(sx_1) = 0,$$

这说明 $\mathrm{Ker}\varphi'' \subseteq \mathrm{Ker}\varphi'$.

现在我们可定义所需要的模同态 φ_1 了. 因为 φ'' 是 $N \oplus R$ 到 N_1 的满同态,又 $\mathrm{Ker}\varphi'' \subseteq \mathrm{Ker}\varphi'$. 我们可合理地定义 N_1 到 W 的模同态 φ_1: $\varphi''(y,r) \mapsto \varphi'(y,r)$,对于 $y \in N$,我们有 $\varphi(y) = \varphi'(y,0) = \varphi_1(\varphi''(y,0)) = \varphi_1(y + 0 \cdot x_1) = \varphi_1(y)$,于是由 φ 可得到 N_1 到 W 的模同态 φ_1.

定理 7.4.9 设 D 是主理想环,M 是有限生成扭模,则 $M = M_1 \oplus \cdots \oplus M_k$,$M_i$ 是阶为 r_i 的循环子模,$r_i \mid r_{i-1}, k \geqslant i \geqslant 2$,$r_1$ 是 M 的指数,进一步地,

$$M \cong D/(r_1) \oplus \cdots \oplus D/(r_k).$$

证明 在M中我们可选取一个元素x_1,其阶为M的指数r_1. 令$M_1 = Rx_1$,又由引理 7.4.6 可知,由恒等映射$\varepsilon: M_1 \to M_1$可得D模同态$\varphi: M \to M_1$. 设$N_2 = \mathrm{Ker}\varphi$,若$x \in M$,因为$\varphi$是由恒等映射扩展得到的,故$\varphi(x - \varphi(x)) = \varphi(x) - \varphi(x) = 0$,从而有$x - \varphi(x) \in N_2$,由此得$x = \varphi(x) + (x - \varphi(x)) \in M_1 + N_2$. 若$x \in M_1 \cap N_2$,则$x = \varphi(x) = 0$,这说明$M = M_1 \oplus N_2$. 对$N_2$继续讨论,因为$M$是有限生成的,所以$N_2$是有限生成的,从而$N_2$有指数,设$N_2$的指数为$r_2$,则$r_2 \mid r_1$,再由引理 7.4.6 知,在$N_2$中可选取元素$x_2$,其阶为$N_2$的指数$r_2$. 令$M_2 = Dx_2$,同上法可证得$N_2 = M_2 \oplus N_3$,因而$M = M_1 \oplus M_2 \oplus N_3$. 这个过程继续下去,因为$M$是有限生成模,$M$满足升链条件,从而必有正整数$m$使$N_m = 0$.

由上面的证法知,对每个i来说,$M_i = Dx_i$. 令$\eta: r \mapsto rx_i$,则η是D到M_i的模同态满射,且$\mathrm{Ker}\eta = (r_i)$,由同态基本定理即得$M_i \cong D/(r_i)$. 定理得证.

定理 7.4.10 设D是主理想环,M是有限生成扭模. 若

$$M = M_1 \oplus \cdots \oplus M_k = N_1 \oplus \cdots \oplus N_t$$

其中M_i是r_i阶的循环模,N_j是S_j阶的循环模,$r_i \mid r_{i-1}, k \geqslant i \geqslant 2, s_j \mid s_{j-1}, k \geqslant j \geqslant 2$,$r_k$与$s_t$都不是$D$的单位,必有$k = t$,且对所有的$i$,$M_i \cong N_i$.

证明 不妨设$k \geqslant t$,在D中选取素元p使得$p \mid r_k$,于是对所有i都有$p \mid r_i$. 因为r_1和s_1都是M的指数,所以r_1与s_1相伴,即$r_1 \sim s_1$,从而$p \mid s_1$,由此可设$p \mid s_j$,p不整除s_{n+1},其中$1 \leqslant j \leqslant n \leqslant t$. 由$p$是素元知,$D/(p) = K$是域. 由引理7.4.7 我们有

$$M(p) = M_1(p) \oplus \cdots \oplus M_k(p) \cong \oplus_{i=1}^{k} K,$$

于是$M(p)$同构于K上的n维向量空间. 又

$$M(p) = N_1(p) \oplus \cdots \oplus N_t(p) = N_1(p) \oplus \cdots \oplus N_n(p) \cong \oplus_{j=1}^{k} K,$$

于是$M(p)$又同构于K上的n维向量空间. 因为$n \leqslant t \leqslant k$,由此得$k = t$,我们已看到$r_1$与$s_l$相伴. 下面证明对所有的$r_i$与$s_i$相伴. 如若不然,设对于$1 \leqslant i \leqslant h$,我们有$r_i$与$s_i$相伴,但$r_h$与$s_h$不相伴. 因为$s_h$不整除$r_h$,令$M' = r_h M$,则

$$M' = r_h M_1 \oplus \cdots \oplus r_h M_{h-1} = r_h N_1 \oplus \cdots \oplus r_h N_{h-1} \oplus r_h N_h \oplus \cdots$$

因为$r_h N_h \neq 0$,这与上面结论相矛盾,于是对所有i来说,$r_i \sim s_i$.

综上所述,对所有i,$M_i \cong D/(r_i) = D/(s_i) \cong N_i$,定理得证.

7.4.3 有限生成模的结构定理

根据定理 7.4.4,7.4.9 及 7.4.10,我们可得主理想环上有限生成模的结构定理.

定理 7.4.11 设D是主理想环,M是有限生成模,则存在$m \in \mathbf{N}^+$及D中非单位元$r_1, \cdots, r_k$使得$M = M_1 \oplus \cdots \oplus M_k \oplus F$,其中$M_i$是阶为$r_i$的循环模,具有关系

$r_i \mid r_{i-1}, k \geqslant i \geqslant 2$，$F$ 是秩为 m 的自由模，则 $M \cong D/(r_1) \oplus \cdots \oplus D/(r_k) \oplus D^m$，即在同构意义下，$M$ 被 m 及 $r_1, \cdots, r_k$ 唯一确定.

定义 7.4.3 R 扭模 M 称为 **p 素模**，若 M 的指数是 R 中某个素元的 p 次幂.

引理 7.4.12 设 M 是主理想环 D 上的有限生成扭模，其指数为 r，若 $r = st$，且 $(s,t) = 1$，则 $M = M(s) \oplus M(t)$.

证明 由题设知，存在 $a, b \in D$ 使得 $as + bt = 1$. 对 $\forall x \in M$，我们有 $x = tbx + sax \in M(s) + M(t)$. 又若 $x \in M(s) \cap M(t)$，则 $x = asx + btx = 0$. 于是有 $M = M(s) \oplus M(t)$.

结合定理 7.4.9，7.4.10，我们可得：

定理 7.4.13 设 M 是主理想环 D 上的有限生成扭模，其指数 $r = \prod\limits_{i=1}^{k} p_i^{e_i}$，其中 p_i 是 D 中两两不相伴的素元，对 $\forall i$ 来说，$e_i \geqslant 1$，则 $M = M_1 \oplus \cdots \oplus M_k$，其中 M_i 为 p_i 素模，且 $M_i = M_{i1} \oplus \cdots \oplus M_{ik}$，其中 $M_{ij} \cong D/(p^{e_{ij}})$ 是阶为 $p^{e_{ij}}$ 的循环模，此时对 $\forall i, j$ 来说，$1 \leqslant e_{ij} \leqslant e_{ij-1} \leqslant e_i$.

在同构意义下，M 由 $\{p_i^{e_{ij}} \mid 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant k_i\}$ 唯一确定.

定义 7.4.4 以上定理中的 $\{p_i^{e_{ij}} \mid 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant k_i\}$ 称为 M 的**初等因子**.

综合定理 7.4.4 及 7.4.13 我们可得有限生成模结构的另一种等价描述.

定理 7.4.14(初等因子型) 设 D 是主理想环，M 是有限生成模，则存在非负整数 m，使得 $M \cong D^m \oplus N$，其中 N 为指数为 r 的有限生成扭模.

若 $r = \prod_{i=1}^{k} p_i^{e_i}$，则 $M \cong \oplus_{i,j} D/(p^{e_{ij}})$，其中 $D/(p^{e_{ij}})$ 是阶为 $p^{e_{ij}}$ 的循环模，对所有的 i, j，$1 \leqslant e_{ij} \leqslant e_{ij-1} \leqslant e_i$. 在同构意义下，$M$ 由 m 及初等因子 $\{p_i^{e_{ij}} \mid 1 \leqslant i \leqslant k, 1 \leqslant j \leqslant k_i\}$ 唯一确定.

若令 $D = \mathbf{Z}$，则可得有限生成交换群结构定理.

例 7.4.2 设 A 是 100 阶的交换群，找出 A 的所有可能类型.

解 按 A 的不变因子型分解，由于 $100 = 50 \times 2 = 20 \times 5 = 10 \times 10$，则 A 有下列四种类型之一：Z_{100}；$Z_{50} \oplus Z_2$；$Z_{20} \oplus Z_5$；$Z_{10} \oplus Z_{10}$.

按 A 的初等因子型分解，由于 $100 = 5^2 \times 2^2$，这里 $p_1 = 5, p_2 = 2$，则 A 的可能初等因子为：$\{5^2, 2^2\}$，$\{5, 5, 2^2\}$，$\{5, 5, 2, 2\}$.

从同构意义上讲，A 可能有下列四种类型之一：

$$Z_{5^2} \oplus Z_{2^2} \cong Z_{100};$$

$$Z_{5^2} \oplus Z_2 \oplus Z_2 \cong (Z_{5^2} \oplus Z_2) \oplus Z_2 \cong Z_{50} Z_2;$$

$$Z_5 \oplus Z_5 \oplus Z_{2^2} \cong Z_5 \oplus (Z_5 \oplus Z_{2^2}) \cong Z_5 \oplus Z_{20};$$

$$Z_5 \oplus Z_5 \oplus Z_2 \oplus Z_2 \cong (Z_5 \oplus Z_2) \oplus (Z_5 \oplus Z_2) \cong Z_{10} \oplus Z_{10}.$$

由此可见,A 的两种分解是一致的.

7.5 有限生成模的自同态环

本节介绍张量积的概念及最基本的结果,它不仅在有限群的表示理论中极其重要,而且在其他代数领域也十分有用,还将讨论如何决定主理想整环上有限生成模 M 的自同态环.

7.5.1 模的张量积

定义 7.5.1 设 R 是有单位元的环,M 是右 R 模,N 是左 R 模. 我们说由 $M\times N$ 到加群 G 的映射 f 是平衡映射,若下列条件满足:

(1) 对于 $\forall m_1,m_2\in M,n\in N$ 都有 $f(m_1+m_2,n)=f(m_1,n)+f(m_2,n)$;

(2) 对于 $\forall m\in M,n_1,n_2\in N$ 都有 $f(m,n_1+n_2)=f(m,n_1)+f(m,n_2)$;

(3) 对于 $\forall m\in M,n\in N,r\in R$ 都有 $f(mr,n)=f(m,rn)$.

定义 7.5.2 M 和 N 在 R 上的张量积是一个加群,记为 $M\odot_R N$,它满足下列条件:有一个平衡映射 η: $M\times N\to M\odot_R N$ 具有性质:

G 是加群,f: $M\times N\to G$ 是一个平衡映射,则存在唯一的群同态 φ: $M\odot_R N\to U$ 使得 $f=\varphi\circ\eta$.

定理 7.5.1 设 R 是有单位元的环,M 是右 R 模,N 是左 R 模,则 M 和 N 在 R 上的张量积存在,且在同构意义下是唯一的.

证明 令 F 是以卡氏积 $M\times N$ 为基的自由 $\mathbf{Z}$ 模,即 F 中元是形如 $n_1(x_1,y_1)+\cdots+n_k(x_k,y_k)$ 的有限和,其中 $(x_i,y_i)\in M\times N,n_i\in\mathbf{Z},1\leqslant i\leqslant k,k\in\mathbf{Z}^+$. F 中任意两个元素相加定义为相同元素的系数相加,令 S 是由下列形状的元素生成的 F 的子加群:

$(m_1+m_2,\ n)-((m_1,\ n)+(m_2,\ n)),(m,\ n_1+n_2)-((m,\ n_1)+(m,\ n_2)),(mr,\ n)-(m,\ rn)$.

$$M\odot_R N=F/S,(m,n)+S=\overline{(m+n)},\forall(m,n)\in F.$$

令 η: $M\times N\to M\odot_R N$,其中,$\eta(m,n)=\overline{(m+n)}$.

于是 η 可看作自然同态 $F\to F/S$ 在 $M\times N$ 上的限制,则在 $M\odot_R N$ 中,我们有下列恒等式:

$$\overline{(m_1+m_2,n)}=\overline{(m_1,n)}+\overline{(m_2,n)};\overline{(m,n_1+n_2)}=\overline{(m,n_1)}+\overline{(m,n_2)};$$

$$\overline{(mr+n)}=\overline{(m+nr)}.$$

于是 η: $M\times N\to M\odot_R N$ 是平衡映射. 现在设 G 是一个加群,f: $M\times N\to G$ 是一

个平衡映射,对任意的$(m,n)\in M\times N$,定义$\varphi(m,n)=f(m,n)$,把φ线性扩张到F,则φ是F到G的同态. 因为f是平衡映射,故$S\subseteq \mathrm{Ker}\varphi$,于是$\varphi$诱导一个映射$\bar{f}$: $M\odot_R N\to G$,其中$\bar{f}(\overline{m+n})=\bar{f}((m,n)+S)=\varphi(m,n)=f(m,n)$,也即$\bar{f}\eta=f$,因为$M\odot_R N$是由所有具有$\overline{(m+n)}$形式的元生成,从而$\bar{f}$是唯一的.

设T_1与T_2分别是M和N关于平衡映射g_1和g_2的张量积,因为T_1是张量积,故存在唯一的群同态h_1: $T_1\to T_2$使得$h_1g_1=g_2$,同理存在唯一的群同态h_2: $T_2\to T_1$使得$h_2g_2=g_1$,由此$(h_1h_2)g_2=g_2$,故h_1h_2是满足$(h_1h_2)g_2=g_2$的T_2到自身的群同态. 又T_2上的恒等自同态I_2显然满足$I_2g_2=g_2$. 由群同态唯一性得$h_1h_2=I_2$. 同理$h_2h_1=I_1$,于是$T_1\cong T_2$,唯一性得证.

因而对$\forall\, m\in M, n\in N$,一般记$m\odot n=\eta(m,n)$.

更具体来说,张量积$M\odot_R N$是由集合$\{m\odot n\mid m\in M, n\in N\}$生成的,其中符号$m\odot n$满足下列恒等式:

1) 对$\forall\, m_1,m_2\in M, n\in N$都有$(m_1+m_2)\odot n=m_1\odot n+m_2\odot n$;

2) 对$\forall\, m\in M, n_1,n_2\in N$都有$m\odot(n_1+n_2)=m\odot n_1+m\odot n_2$;

3) 对$\forall\, m\in M, n\in N, r\in R$都有$(mr)\odot n=m\odot(rn)$.

注意,$M\odot_R N$的元素并不完全是符号$m\odot n$,而是这样一些符号的和.

由张量积的定义知,张量积是一个加群,它没有模结构. 然而在适当条件下,张量积可以是个模.

定理 7.5.2　设R,S是有单位元的环,M是(S,R)双模,N是左R模. 若规定$s(m\odot n)=(sm)\odot n$, $s\in S$, $n\in N, m\in M$,则$M\odot_R N$是左S模.

证明　对于$(m,n)\in M\times N$,令$\eta(m,n)=(sm)\odot n$,则η是$M\times N$到$M\odot_R N$的一个平衡映射,因而存在$M\odot_R N$到自身的同态μ_s使得$\mu_s(m\odot n)=(sm)\odot n$. 规定$s(m\odot n)=\mu(m\odot n)$,则$M\odot_R N$是一个左$S$模.

我们知道,若R是交换环,则任何一个R模都可作成一个(R,R)双模. 为了更好地理解张量积,我们给出一个具体例子.

例 7.7.1　设F是域,U和V是F上向量空间,则U是一个(F,F)双模,因而我们可以定义向量空间$U\odot_F V$. 设$\{u_1,\cdots,u_r\}$与$\{v_1,\cdots,v_s\}$分别是U和V的基,则$U\odot_F V$是一个rs维的向量空间,其中$\{u_i\odot v_j\mid 1\leqslant i\leqslant r, 1\leqslant j\leqslant s\}$是$U\odot_F V$的一个基.

若$u=\sum\limits_i a_iu_i\in U, v=\sum\limits_j b_jv_j\in V$,则$u\odot v=\sum\limits_{i,j}a_ib_j(u_i\odot v_j)$.

定理 7.5.3　设R是有单位元的环,M是R模,则M与$R\odot_R M$同构.

证明　对于$r\in R, m\in M$,令$f(r,m)=rm$,则f是$R\times M$到M的映射,因而诱导出一个R模同态η: $R\odot_R M\to M$,其中$\eta(r\odot m)=rm$. 令ξ: $m\to 1\odot m$为$M\to R\odot_R M$的模同态,则$\eta\xi=I$,因而η是同构.

定理 7.5.4 设 R 和 S 是有单位元的环，$M_1,\cdots,M_r$ 是 (R,S) 双模，N 是 S 模，则 $(\bigoplus_i M_i)\odot_S N$ 与 $\bigoplus_i(M_i\odot_S N)$ 作为 R 模是同构的.

证明 定义

$f:(\bigoplus_i M_i)\times N\to\bigoplus_i(M_i\odot_S N)$，$f((m_1,\cdots,m_r),n)=(m_1\odot n,\cdots,m_r\odot n)$，

则 f 是平衡映射，因而诱导出一个由 $(\bigoplus_i M_i)\odot_S N$ 到 $\bigoplus_i(M_i\odot_S N)$ 的同态 η. 类似地，我们可定义由 $\bigoplus_i(M_i\odot_S N)$ 到 $(\bigoplus_i M_i)\odot_S N$ 的逆映射 ξ 使得 $\eta\xi=I$. 所以 η 是同构.

定理 7.5.5 设 F 是域，U 和 V 是 F 上向量空间，其中 $\dim_F U<\infty$，令 $U^*=\mathrm{Hom}_F(U,F)$. 定义 ξ: $U^*\odot_F V\to\mathrm{Hom}_F(U,V)$，$\xi(\varphi\odot v)(u)=\varphi(u)v$，作线性扩张，其中 $\varphi\in U^*$，$v\in V$，$u\in U$，则作为 F 向量空间，$U^*\odot_F V\cong\mathrm{Hom}_F(U,V)$.

证明 因为 ξ 是由平衡映射 $(\varphi,u)\in U^*\times V\to\varphi(u)v\in V$ 诱导的同态. 令 $\{u_1,\cdots,u_n\}$ 是 U 的一个基，对 $\forall 1\leqslant i\leqslant n$，令 $f_i\in U^*$，对任意的 j，$f_i(u_j)=\delta_{ij}$，则 $\{f_1,\cdots,f_n\}$ 是 U^* 的一个基. $U^*\odot_F V$ 的每个元 $\sum_k(\sum_i\alpha_{ik}f_i)\odot v_k$ 可以写为 $\sum_i f_i\odot(\sum_k\alpha_{ik}v_k)$，因而对某个 $v_i\in V$ 来说，$U^*\odot_F V$ 的每个元都具有 $\sum_{i=1}^n f_i\odot v_i$ 的形式. 取 $\sum_{i=1}^n f_iv_i\in U^*\odot_F V$，固定某个 j，则

$$\xi\Big(\sum_{i=1}^n f_i\odot v_i\Big)(u_j)=\sum_{i=1}^n\xi(f_i\odot v_i)(u_j)=\sum_{i=1}^n f_i(u_j)=v_j.$$

于是，如果 $\xi\big(\sum_{i=1}^n f_i\odot v_i\big)=0$，对所有 j，$v_j=0$，则 $\sum_{i=1}^n f_i\odot v_j=0$. 这就说明 ξ 是单同态.

现在令 $\sigma\in\mathrm{Hom}_F(U,V)$，则对于 $u=\sum_{i=1}^n\alpha_iu_i\in U$，我们有

$$\xi\Big(\sum_{i=1}^n f_i\odot\sigma(u_i)\Big)(u)=\sum_{i=1}^n f_i(u)\,\sigma(u_i)=\sum_{i=1}^n\alpha_i\sigma(u_i)=\sigma(u).$$

这说明 $\xi\big(\sum_{i=1}^n f_i\odot\sigma(u_i)\big)=\sigma$，因而 ξ 是满射.

7.5.2 主理想整环上有限生成模的自同态环

接下来我们讨论如何决定主理想整环上有限生成模 M 的自同态环 D'(即 $\mathrm{Hom}(M,M)$. 从 M 的一个分解开始：$M=Dz_1\oplus Dz_2\oplus\cdots\oplus Dz_s$，

$$a_mz_1\Leftrightarrow a_mz_2\Leftrightarrow\cdots a_mz_s=(d_i)\neq0,i\leqslant r,$$

但当 $i>r$ 时 $a_mz_i=0$. 设 $\eta\in D'$，并假设 $\eta z_i=\eta(z_i)=w_i\in M$，$1\leqslant i\leqslant s$，那么，如果 $x\in M$，$x=\sum_1^s a_iz_i$，$a_i\in D$，因此 $\eta x=\eta(\sum a_iz_i)=\sum\eta(a_iz_i)=\sum a_i(\eta z_i)=$

$\sum a_i w_i$，这表明 η 是由它在 M 的生成元 z_i 上的作用所决定的，且 $d_i w_i = d_i(\eta z_i) = \eta(d_i z_i) = 0$，这表明 $a_m w_i \Leftrightarrow a_m z_i$. 所以，如果 $a_m w_i = (g_i)$，则当 $i > j$ 时 g_i 是任意的，而当 $i \leqslant r$ 时 $g_i \mid d_i$.

反之，假设对每个 i，挑选一个元 $w_i \in M$ 使 $a_m w_i \supset a_m z_i$，假设 $x \in M$，而 $x = \sum_1^s a_i z_i = \sum b_i z_i$ 是 x 的两个表示形式，那么我们有 $a_i - b_i \in a_m z_i$，因此，$a_i - b_i \in a_m w_i$，从而有结果 $\sum a_i w_i = \sum b_i w_i$. 这表示 η: $\sum a_i z_i \to \sum a_i w_i$ 是 M 到 M 的映射. 直接验证也证明 $\eta \in D'$.

从环 $D = \mathrm{Hom}(M,M)$ 到 M 的 s -元序组的集合上的满足 $a_m w_i \supset a_m z_i$ 的双射

$$\eta \to (w_1, w_2, \cdots, w_s).$$

令 $w_i = \sum b_{ij} z_j, b_{ij} \in D$，对这个有序组 $(w_1, w_2, \cdots, w_s)$ 结合一个元在 D 里 $s \times s$ 矩阵环 $M_s(D)$ 里的矩阵 $(b_{ij})_{ss}$，这个矩阵可以不是唯一决定的. 因为任意的 b_{ij} 可以用 b'_{ij} 代替，在 $j \leqslant r$ 时保持 $b'_{ij} \equiv b_{ij} (\mathrm{mod} d_{ij})$，这是保持 w_i 不变的仅有的变动，条件 $a_m w_i \supset a_m z_i$ 等价于 $d_i b_{ij} \equiv 0 (\mathrm{mod} d_j)$，这自然是指存在 $c_{ij} \in D$ 使 $d_i b_{ij} = c_{ij} d_j$，因此存在一个 $C \in M_s(D)$ 使

$$\mathrm{diag}\{d_1, d_2, \cdots, d_s\} B = C\mathrm{diag}\{d_1, d_2, \cdots, d_s\},$$

适合上式的矩阵 B 的集合 R 是 $M_s(D)$ 的一个子环. 任意 $B \in R$ 决定一个 $\eta \in D'$ 使 $\eta z_i = \sum b_{ij} z_j$. 容易验证映射 $B \to \eta$ 是从 R 到 D' 上的一个满同态. 显然 $\eta = 0$，当且仅当对 $B = (b_{ij})$ 有 $b_{ij} \equiv 0 (\mathrm{mod} d_j)$，因而同态的核 k 是使 $B = Q\mathrm{diag}\{d_1, d_2, \cdots, d_s\}$ 的矩阵 B 的集，这里 $Q \in M_s(D)$. 我们指出，这种形式的矩阵自动地满足 $\mathrm{diag}\{d_1, d_2, \cdots, d_s\} B = C\mathrm{diag}\{d_1, d_2, \cdots, d_s\}$.

定理 7.5.6 设 $M = Dz_1 \oplus Dz_2 \oplus \cdots \oplus Dz_s$，这里阶理想 $a_m z_i = (d_i)$ 适合

$$a_m z_1 \supset a_m z_2 \supset \cdots \supset a_m z_s,$$

于是 D-模 M 的自同态环 D' 与 R/k 反同构，这里 B 是由矩阵 $B \in M_s(D)$ 所组成的环，对于 B 有一个 $C \in M_s(D)$ 使

$$\mathrm{diag}\{d_1, d_2, \cdots, d_s\} B = C\mathrm{diag}\{d_1, d_2, \cdots, d_s\},$$

k 是形如 $Q\mathrm{diag}\{d_1, d_2, \cdots, d_s\}$ 的矩阵的理想，$Q \in M_s(D)$.

如果 M 是自由模，所有 $d_i = 0$，于是 $B \in M_s(D)$ 及 $k = 0$. 在这种情况下，如果 $s = 1$，M 是循环的，根据 D 的可换条件 $B = (b_{ij})$ 当然是适合的，于是 $D' \cong D/(d)$，这里 $d = d_i$. 矩阵环 R 是可以确定的.

假如我们使用 d_i 上的条件 $d_i \mid d_j$(当 $i \leqslant j \leqslant r$ 时) 和 $d_i = 0$(当 $i > r$ 时)，以及

$d_i b_{ij} \equiv 0 (\bmod d_j)$，则：

1）如果 $i \geqslant j$，那么 b_{ij} 是任意的，因为此时 $d_i \equiv 0(\bmod d_j)$；

2）如果 $i \leqslant r, j > r$，那么 $b_{ij} \equiv 0$，因为此时 $d_i \neq 0$ 和 $d_j = 0$；

3）如果 $i, j > r$，那么 b_{ij} 是任意的，因为此时 $d_i = d_j = 0$；

4）如果 $i < j \leqslant r, b_{ij} \equiv 0(\bmod d_i^{-1} d_j)$ 略变动符号，B 将如下式

$$\begin{pmatrix} b_{11} & b_{12} d_1^{-1} d_2 & \cdots & b_{1r} d_1^{-1} d & 0 & \cdots & 0 \\ b_{21} & b_{22} & \cdots & b_{2r} d_1^{-1} d & 0 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ b_{r1} & b_{r2} & \cdots & b_{rr} & 0 & \cdots & 0 \\ b_{r+1,1} & b_{r+1,2} & \cdots & b_{r+1,r} & b_{r+1,r+1} & \cdots & b_{r+1,s} \\ b_{s1} & b_{s2} & \cdots & b_{sr} & b_{s,r+1} & \cdots & b_{ss} \end{pmatrix}$$

这里右上角全为0，凡表示为 b_{ij} 的都是任意的，当 $i < j \leqslant r$ 时，(i,j) 处的元素 $b_{ij} d_i^{-1} d_j$ 属于矩阵的条件是，当 $i \geqslant j$ 和 $j \leqslant r$ 时，b_{ij} 被 d_j 除尽；当 $i < j \leqslant r$ 时，b_{ij} 被 d_i 除尽，如果这模是挠模，$r = s$ 以及上面矩阵简化为左上角的矩阵块.

特别地，我们考虑 $M = V$ 的情况，这里 V 是由 F 上的有限维向量空间 V 的线性变换 T 所决定的 $F[\lambda]$-模，这是一个挠模，任意 $b_{ij}, i \geqslant j$，可以被模为 d_j 的同一陪集里的 b'_{ij} 代替. 因此，我们可以假定 $\deg b_{ij} < n_j = \deg d_j$，当 $i \geqslant j$ 时. 同样，我们假定 $\deg b_{ij} < n_i$，当 $i < j$ 时. 适合这些条件的矩阵 $B \in R$ 叫正规化的，限制在 R 的正规化矩阵的映射 $B \to \eta$ 是到 D' 里的一个双射，有一个自然的方式把 D' 和 R 当作 F 上的向量空间，只要用 $a \in F$ 乘 $B \in R$ 的所有元，对 R 就构成 F 上的模，对 D' 我们用 $(a\eta)x = a(\eta)x = \eta(ax)$ 来定义，这里 $a \in F, \eta \in D$，在 R 里的正规化矩阵的集合 S 是一个子空间，$B \to \eta$ 是从 S 到 D' 里的一个 F-线性同构.

设 $S_{ij}, 1 < i, j \leqslant s$，表示除 (i, j) 一元外其他的元都是0的正规化矩阵所组成的 S 的子空间，那么 $\dim S_{ij} = n_j$，当 $i \geqslant j$ 以及 $\dim S_{ij} = n_j$ 当 $i < j$. 因为 S 是子空间 S_{ij} 的直和，所以有

$$\dim S = \sum_{j=1}^{s} (s - j + 1) n_j + \sum_{i=1}^{s-1} (s - i) n_i = \sum_{j=1}^{s} (2s - 2j + 1) n_j.$$

定理 7.5.7 (Frobenius)　设 F 是域，$A \in F_n$，又设 $d_1(\lambda), d_2(\lambda), \cdots, d_s(\lambda)$ 是 $\lambda I - A$ 里不等于 I 的不变因子，再设 $n_i = \deg d_i(\lambda)$，那么与 A 可交换的矩阵组成的 F 上的向量空间的维数是由下面的公式给出的：

$$N = \sum_{j=1}^{s} (2s - 2j + 1) n_j.$$

当然也能用线性变换来叙述如下的推论：一个线性变换 T 是循环的（即对应的

$F[\lambda]$-模是循环的)，当且仅当与 T 可交换的线性变换是 T 的多项式.

证明　T 是循环的当且仅当 $s=1$，我们也知道 $d_s(\lambda)$ 是 T 的最小多项式 $m(\lambda)$，因此 n_s 是系数在 F 上的 T 的多项式环 $F[T]$ 的维数，如果 $s=1$，那么 $N=n_1=\dim F[T]$ 与 T 可交换的线性变换的空间当然是包含 $F[T]$ 的，而且与 $F[T]$ 一致. 如果 $s>1$，那么

$$N=\sum_{j=1}^{s}(2s-2j+1)n_j,$$

即 $N>\sum n_j>n_s$，因此存在非 T 的多项式而与 T 可换的线性变换.

定理 7.5.8　$D'=\mathrm{Hom}(M,M)$ 的中心是映射 $x\to ax$，$a\in D$ 的集合.

证明　D' 的决定说明对 $\forall i, 1\leqslant i\leqslant s$，存在一个自同态 ε_{is} 使 $\varepsilon_{is}z_s=z_i$，$\varepsilon_{is}z_j=0$，当 $j\neq s$ 时. 现在设 r 在 D' 的中心里，那么

$$rz_s=r\varepsilon_{ss}z_s=\varepsilon_{ss}rz_s=\varepsilon_{ss}\left(\sum a_iz_i\right)=\sum a_i\varepsilon_{ss}z_i=a_sz_s, a_i\in D,$$

并且

$$rz_i=r\varepsilon_{is}z_s=\varepsilon_{is}rz_s=\varepsilon_{ss}\left(\sum a_jz_j\right)=\sum a_i\varepsilon_{ss}rz_j=a_sz_i,$$

从而 r 是映射 $x\to ax$.

推论 7.5.9　如果 U 是一个有限维的向量空间里的一个线性变换，它与每一个与已给的线性变换 T 交换的线性变换都可交换，那么 U 是 T 的一个多项式.

取 $T=1$ 时可以得到下面的结果：

推论 7.5.10　域 F 上有限维向量空间的线性变换环的中心是数乘法 $x\to ax$，$a\in F$ 的集合.

第 8 章　向量空间的分解和算子的若当标准型

在本章中，我们将讨论有限维向量空间的分解和线性算子的若当标准型问题，首先运用上一章的模分解定理来研究线性算子的结构，然后讨论向量空间的分解、有理标准型、若当标准型及射影代数等问题. 本章中的向量空间都假定为是有限维的.

8.1　带有线性算子的模

8.1.1　线性算子

有限维向量空间上的任何线性算子都可以用矩阵乘法来表示.

定理 8.1.1　令 $\tau \in L(V)$，$B = (b_1, \cdots, b_n)$ 是 V 的一个有序基，那么 τ 可以用线性算子 $\tau_A \in L(F^n)$ 表示，即 $(\tau(\nu))_B = \tau_A((\nu)_B)$. 这里 $A = (\nu)_B$ 是第 i 列为 $(\tau(\nu))_B$ 的矩阵，$(\tau(\nu))_B = (\tau)_B(\nu)_B$.

因为矩阵 $(\tau)_B$ 与有序基 B 的选取有关，所以我们自然想知道如何选取基 B，以使矩阵 $(\tau)_B$ 尽可能简单.

同样，我们重新叙述一下 τ 的与不同有序基相关的矩阵间的关系.

定理 8.1.2　令 $\tau \in L(V)$，B 和 B' 是 V 的有序基，那么 τ 关于 B' 的矩阵可以用 τ 关于 B 的矩阵表示成 $(\tau)_B = M(\tau)_B M^{-1}$，其中 M 是基矩阵的变换，其第 i 列为 $(b_i)_{B'}$，这里 $B = (b_1, \cdots, b_n)$.

定义 8.1.1　两个矩阵 S，T 是**相似的**，如果存在一可逆矩阵 P，使得 $T = PSP^{-1}$. 带有相似性的等价类叫做**相似类**.

定理 8.1.3　对于矩阵 S 和 T，以下论断是等价的：

1) 如果 S 表示关于有序基 B 的线性算子 $\tau: V \to V$，那么 T 也表示 τ，但关于另一个不同的有序基，即，如果 $S = (\tau)_B$，那么存在有序基 B'，使得 $B = (\tau)_B$.

2) S 与 T 相似.

根据定理 8.1.3，表示给定的线性算子 $\tau \in L(V)$ 的矩阵正好是特殊相似类中的矩阵. 因此，为了能最好地表示出 τ，我们要找出这一相似类的简单代表. 更一般地，为了表示 V 上所有的线性算子，我们需要找出每个相似类的简单代表，即，相似的简单的典

范型集.

例 8.1.1 V 的子集 W 是 $F[x]$模的子模$\Leftrightarrow W$ 是向量空间 V 的一个不变子空间.

例 8.1.2 $F[x]$中的单位正好是 F 中的非零标量.

定义 8.1.2 令 $\sim$ 是 W 上一等价关系. 子集 $K \subset W$ 称为 $\sim$ 的**典范型集**,如果对于每一个 $s \in W$,恰存在 $k \in K$,使得 $k \sim s$.

目前,最简单的矩阵可能是对角阵. 但并不是所有的线性算子都可以用对角阵表示. 换句话说,对角阵集并不构成相似的典范型集.

这就为进一步研究提供了两个不同的方向.

一是对于那些可以由对角阵表示的线性算子,我们可以找出它们的特征. 这种算子称为可对角化算子.

二是可以找出一种不同类型的"简单"矩阵,它能给出相似的典范型集.

我们将同时沿着这两个方向继续讨论.

我们关注的是非零线性算子 $\tau \in L(V)$,而 V 不仅是域 F 上的向量空间,同时也是 $F[x]$ 上的一个模,有由 $p(x)v = p(\tau)(v)$ 定义的数乘. 我们准备通过联系模和向量空间的概念,把上一章的语言转化为 V 的语言.

首先,因为 V 是一有限维向量空间,所以模 V 是挠模. 为了说明这一点,观察到同构于 $M_n(F)$ 的向量空间 $L(V)$ 有维数 n^2. 因此,n^2+1 个向量 $\iota, \tau, \cdots, \tau^2, \tau^{n^2}$ 线性相关,则对于任意多项式 $p(x) \in F[x]$,$p(\tau) = 0$. 因此,$p(x)V = \{0\}$,所以 V 的所有元素都是挠元.

同样,作为一个模,V 是有限生成的. 因为如果 $B = \{v_1, \cdots, v_n\}$ 是向量空间 V 的一个基,那么每个向量 $v \in V$ 都是一个线性组合 $v = r_1v_1 + \cdots + r_nv_n$,其中 $r_i \in F \subset F[x]$,所以 B 生成了模 V.

因此,V 是主理想整环 $F[x]$ 上一有限生成挠模,可以运用上一章的分解定理.

8.1.2 极小多项式

在模 V 的子模和向量空间 V 的子空间之间有一种简单的联系. 如果 $\tau(W) \subset W$,那么 V 的子空间 W 在 τ 之下是不变的.

定理 8.1.4 V 的子集 W 是 $F[x]$模 V 的一个子模,当且仅当 W 是向量空间 V 的不变子空间.

由此定理知,V 的子模 W 转化为 $F[x]$ 模有两种方法 —— 作为 V 的一个子模;作为一个运用 τ 到 S 的限制 $\tau|_W: W \to W$ 的模. 但是,因为对于 $\forall s \in W$,$p(\tau)(s) = p(\tau|_W)(s)$,数乘在这两种情形下是一样的,所以这两个模是一样的.

下面考虑 V 的零化子 $\mathrm{ann}(V) = \{p(x) \in F(x) \mid p(x)V = \{0\}\}$. 它是 $F(x)$ 的非零主理想. 因为 V 的所有阶(即 $\mathrm{ann}(V)$ 的生成元)都是相伴的,又因为 $F(x)$ 的单位正

好是 F 的非零元，所以存在 V 的唯一的首一阶. 这就有了以下定义：

定义 8.1.3　模 V 的唯一的首一阶，即生成 $\mathrm{ann}(V)$ 的唯一的首一多项式，叫做 τ 的**极小多项式**. 我们用 $m_\tau(x)$ 或 $\min(\tau)$ 表示这一多项式. 因此，$\mathrm{ann}(V) = < m_\tau(x) >$；$p(x)V = \{0\}$，当且仅当 $m_\tau(x) \mid p(x)$；或等价地，$ap(\tau) = 0$，当且仅当 $m_\tau(x) \mid p(x)$.

我们可以简单地把线性算子 τ 的极小多项式定义为：唯一地使得 $m_\tau(\tau) = 0$ 的含有最小次数的首一多项式. 阶与极小多项式之间的联系也传给了子模.

定理 8.1.5　令 W 是模 V 的子模，那么 W 的首一阶是限制 $\tau|_W$ 的极小多项式.

证明　已知如果 $q(x)$ 是 W 的首一阶，那么

$$< q(x) > = \mathrm{ann}(W) = \{ p(x) \mid p(x)W = \{0\} \} = \{ p(x) \mid p(\tau|_W)(W) = \{0\} \}.$$

所以 $q(x)$ 是限制 $\tau|_W$ 的极小多项式.

极小多项式的概念也可以定义在矩阵上. 特别地，如果 S 是 F 上一方阵，那么 S 的极小多项式 $m_S(x)$ 是唯一的使得 $p(S) = 0$ 的含有最小次数的首一多项式 $p(x) \in F(x)$. 而且以上概念是唯一确定的 并且下面的定理也成立：

定理 8.1.6　1) 如果 S, T 是相似阵，那么 $m_S(x) = m_T(x)$. 因此，极小多项式在相似性之下是一个不变式.

2) $\tau \in L(V)$ 的极小多项式与任意表示 τ 的矩阵的极小多项式相同.

8.1.3　循环子模与循环子空间

考虑循 环子模 $< v > = \{p(x)v \mid p(x) \in F(x)\}$，并假设它有首一阶 $m(x)$. 因此，$m(x)$ 是限制 $\sigma = \tau|_{<v>}$ 的极小多项式. 如果 $m(x) = a_0 + a_1 x + \cdots + a_{n-1}x^{n-1} + x^n$，那么 $B = (v, xv, \cdots, x^{n-1}v) = (v, \sigma(v), \cdots, \sigma^{n-1}(v))$ 是向量空间 $< v >$ 的有序基. 为了说明 B 是线性无关的，假设存在非零标量，使得 $r_0 v + r_1 xv + \cdots + r_{n-1}x^{n-1}v = 0$，即 $(r_0 + r_1 x + \cdots + r_{n-1}x^{n-1})v = 0$，那么 $(r_0 + r_1 x + \cdots + r_{n-1}x^{n-1}) < v > = \{0\}$，所以 $m(x) \mid r_0 + r_1 x + \cdots + r_{n-1}x^{n-1}$，则对于 $\forall i = 1, \cdots, n-1, r_i = 0$.

对于 任意多项式 $p(x) \in F[x]$，$< v >$ 的所有元素都有形式 $p(x)v$. 用极小多项式 $m(x)$ 除 $p(x)$，得 $p(x) = q(x)m(x) + r(x)$，这里 $\deg r(x) < \deg m(x)$. 由于 $m(x)v = 0$，所以我们有 $p(x)v = q(x)m(x)v + r(x)v = r(x)v$. 这表明 $< v >$ 的所有元素都有形式 $r(x)v$，其中 $\deg r(x) < \deg q(x) = n$，用符号记为 $< v > = \{r(x)v \mid \deg r(x) < \deg m(x)\}$. 因此，如果 $r(x) = r_0 + r_1 x + \cdots + r_{n-1}x^{n-1}$，我们就有 $r(x)v = r_0 v + r_1 xv + \cdots + r_{n-1}x^{n-1}v \in \mathrm{span}(B)$，从而 B 是 $< v >$ 的一个有序基.

为了确定 σ 关于 B 的矩阵 $(\sigma)_B$，对于 $i = 0, \cdots, n-2$，$\sigma(\sigma^i(v)) = \sigma^{i+1}(v)$. 所以 σ 只

不过是把 B 中的每个基向量“移到”下一个基向量中. 同样, $m(\sigma)=0$, 则

$$\sigma(\sigma^{n-1}(v))=\sigma^{n}(v)=-(a_0+a_1\sigma+\cdots+a_{n-1}\sigma^{n-1})(v)$$
$$=-a_0v-a_1\sigma(v)-\cdots-a_{n-1}\sigma^{n-1}(v).$$

定义 8.1.4　σ 关于 B 的矩阵为

$$U(m(x))=\begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & 0 & -a_{n-2} \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

该矩阵称为多项式 $m(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}+x^n$ 的**友矩阵**.

例 8.1.3　若 U 是 $p(x)$ 的友矩阵, 那么 $p(U)=0$, 且 U 有极小多项式 $p(x)$.

定义 8.1.5　令 $\tau\in L(V)$, 称 V 的子空间 W 是 **τ 循环的**, 如果存在向量 $v\in W$, 使得集合 $\{v,\tau(v),\cdots,\tau^{m-1}(v)\}$ 是 W 的一个基, 这里 $m=\dim(W)$.

定理 8.1.7　1) 子集 $W\subset V$ 是 V 的一循环子模 $\Leftrightarrow$ 它是 V 的一个 τ 循环子空间.

2) 设 $\langle v\rangle$ 是 V 的一循环子模. 如果 $\langle v\rangle$ 的首一阶($\sigma=\tau|_{\langle v\rangle}$ 的极小多项式) 为

$$m_\sigma(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}+x^n,$$

那么 $B=(v,xv,\cdots,x^{n-1}v)=(v,\sigma(v),\cdots,\sigma^{n-1}(v))$ 是 $\langle v\rangle$ 的一个有序基, 并且 $(\sigma)_B$ 是 $m_\sigma(x)$ 的友矩阵 $U(m(x))$.

8.1.4　向量空间的循环分解定理

下表总结了模概念与向量空间概念之间的关系.

模 V	向量空间 V
数乘: $p(x)v$ 子模 零算子: $\mathrm{ann}(v)=\{p(x)\mid p(x)V=\{0\}\}$ V 的首一阶: $\mathrm{ann}(V)=\langle m(x)\rangle$ 循环子模: $\langle v\rangle=\{p(x)v\mid p(x)\in F[x]\}$	τ 的作用: $p(\tau)(v)$ 不变子空间 零算子: $\mathrm{ann}(V)=\{p(x)\mid p(\tau)(V)=\{0\}\}$ τ 的极小多项式: $m(x)$是使得 $m(\tau)=0$ 的含有最小次数的多项式 τ 循环子空间: $\langle v\rangle=\mathrm{span}\{v,\tau(v),\cdots,\tau^{m-1}(v)\}$

现在, 我们可以把循环分解定理转化为向量空间的语言了.

定理 8.1.8(向量空间的循环分解定理)　令 $\tau \in (V)$，其中 $\dim(V) < \infty$. 如果 τ 的极小多项式为 $m_\tau(x) = p_1^{e_1}(x) \cdots p_n^{e_n}(x)$，这里首一多项式 $p_i(x)$ 是互异和不可约的，那么 V 为直和

$$V = V_{p_1} \oplus \cdots \oplus V_{p_n},$$

其中 $V_{p_i} = \{ v \in V \mid p_i^{e_i}(\tau)(v) = 0 \}$ 是 V 的不变子空间(子模)，且

$$\min(\tau|_{V_{p_i}}) = p_i^{e_i}(x).$$

此外，每个 V_{p_i} 都可以进一步分解为 τ 循环子空间(循环子模)的直和

$$V_{p_i} = < v_{i,1} > \oplus \cdots \oplus < v_{i,k_i} >,$$

这里

$$\min(\tau|_{< v_{i,j} >}) = p_i^{e_{i,j}}(x), e_i = e_{i,1} \geqslant e_{i,2} \geqslant \cdots \geqslant e_{i,k_i}.$$

V 的初等因子 $p_i^{e_{i,j}}(x)$ 即 τ 的初等因子，由算子 τ 唯一确定.

这样 V 的分解就变为 τ 循环子空间的一个直和的分解：

$$V = (< v_{1,1} > \oplus \cdots \oplus < v_{1,k_1} >) \oplus \cdots \oplus (< v_{n,1} > \oplus \cdots \oplus < v_{n,k_n} >).$$

定理 8.1.9　$\tau \in L(V)$ 的极小多项式是它的初等因子的最小公倍式.

8.2　有理典范型

8.2.1　分块对角形矩阵

我们可以用循环分解定理来确定相似的典范型集.

定义 8.2.1　如果 $V = S \oplus T$，且 S 和 T 在 τ 下都是不变的，那么称 **(S, T) 约化 τ**.

如果限制 $\tau|_S: S \to S, \tau|_T: T \to T$ 分别为 S 和 T 上的线性算子，(S, T) 约化 τ. 如果存在 V 的子空间 S 和 T，使 (S, T) 约化 τ，且 $\rho = \tau|_S, \sigma = \tau|_T$，记为 $\tau = \rho \oplus \sigma$.

如果 $\tau = \rho \oplus \sigma$，那么 ρ 与 σ 的任意矩阵表示都可以用来构造 τ 的矩阵表示. 这与我们所讨论的情形特别相关，因为根据定理 8.1.8，

$$\tau = \tau|_{< v_{1,1} >} \oplus \cdots \oplus \tau|_{< v_{n,n} >}. \tag{8.1}$$

我们引入分块矩阵的符号表示法. 如果 A 是 $n \times n$ 矩阵，B 是 $m \times m$ 矩阵，那么由分块矩阵的表示法有

$$M = \begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}.$$

定理 8.2.1　设 $\tau=\tau_1\oplus\tau_2\in L(V)$，其对应的约化对为 (S, T). 令 $X=(c_1, \cdots, c_s)$ 是 S 的一个有序基，$\Delta=(d_1, \cdots, d_t)$ 是 T 的一个有序基，且令 $B=(c_1, \cdots, c_s, d_1, \cdots, d_t)$ 是与 V 相应的有序基，那么矩阵 $(\tau)_{B'}$ 有分块对角型

$$(\tau)_{B'}=\begin{bmatrix}(\tau_1)_C & 0\\ 0 & (\tau_2)_D\end{bmatrix}.$$

上面 (8.1) 式中，如果 $B_{i,j}$ 是循环子模 $<v_{i,j}>$ 的一个有序基，且若 $B=(B_{1,1}, \cdots, B_{n,k_n})$ 表示从这些有序基中得到的 V 的有序基，那么

$$(\tau)_{B'}=\begin{bmatrix}(\tau_{1,1})_{\beta_{1,1}} & & \\ & \ddots & \\ & & (\tau_{n,k_n})_{\beta_{n,k_n}}\end{bmatrix},$$

其中 $\tau_{i,j}=\tau|_{<i,j>}$.

循环子模 $<v_{i,j}>$ 有首一阶 $p_i^{e_{i,j}}(x)$，即限制 $\tau_{i,j}$ 有极小多项式 $p_i^{e_{i,j}}(x)$. 因此若 $\deg p_i^{e_{i,j}}(x)=d_{i,j}$，那么 $B_{i,j}=(v_{i,j}, \tau_{i,j}(v_{i,j}), \cdots, \tau_{i,j}{}^{d_{i,j}-1}(v_{i,j}))$ 是 $<v_{i,j}>$ 的一个有序基.

8.2.2　有理典范型

从而，我们就有了下面定理中描述的 τ 的矩阵表示法.

定理 8.2.2　令 $\dim(V)<\infty$，设 $\tau\in L(V)$ 含有极小多项式

$$m_\tau(x)=p_1^{e_1}(x)\cdots p_n^{e_n}(x),$$

这里首一多项式 $p_i(x)$ 是互异和不可约的，那么我们可以记为

$$V=(<v_{1,1}>\oplus\cdots\oplus<v_{1,k_1}>)\oplus\cdots\oplus(<v_{n,1}>\oplus\cdots\oplus<v_{n,k_n}>),$$

其中 $<v_{i,j}>$ 是 V 的 τ 循环子空间. $\tau_{i,j}=\tau|_{<v_{i,j}>}$ 的极小多项式为 V 的初等因子

$$\min(\tau_{i,j})=p_i^{e_{i,j}}(x),$$

其中 $e_i=e_{i,1}\geqslant e_{i,2}\geqslant\cdots\geqslant e_{i,k_i}$.

这些初等因子由 τ 唯一确定. 而且，如果 $\deg p_i^{e_{i,j}}(x)=d_{i,j}$，那么

$$B_{i,j}=(v_{i,j}, \tau_{i,j}(v_{i,j}), \cdots, \tau_{i,j}{}^{d_{i,j}-1}(v_{i,j}))$$

是 $<v_{i,j}>$ 的一个有序基，且 τ 的关于有序基 $B=(B_{1,1}, \cdots, B_{n,k_n})$ 的矩阵为分块对角阵

$$(\tau)_{B'}=\begin{bmatrix} C(p_1^{e_{1,1}}(x)) & & & & & & \\ & \ddots & & & & & \\ & & C(p_1^{e_{1,k_1}}(x)) & & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & C(p_n^{e_{n,n_1}}(x)) & \\ & & & & & & \ddots \\ & & & & & & & C(p_n^{e_{n,k_n}}(x)) \end{bmatrix}$$

称为 τ 的**有理典范型**.

我们用 $\mathrm{diag}(C(p_1^{e_{1,1}}(x)),\cdots,C(p_n^{e_{n,k_n}}(x)))$表示右边的矩阵.

由定理 8.2.2 可得,对于任意 $\tau\in L(V)$,我们都可以找到一个有序基 B,使得矩阵 $(\tau)_{B'}$ 含有有理典范型.

另一方面,τ 仅有一种有理典范型(最多是重新排列对角线上的分块). 为了说明这一点,假设对于 V 的某个有序基 X,矩阵 $(\tau)_X$ 有形式

$$(\tau)Z_X=\mathrm{diag}(C(q_1^{f_{1,1}}(x)),\cdots,C(q_m^{f_{m,k_m}}(x))),$$

那么 V 是含有初等因子 $q_i^{f_{i,j}}(x)$的循环子模的一个直和. 因此,有理典范型的唯一性(最多是重排对角线上的分块)是从 V 的循环分解的唯一性中得出的.

定理 8.1.10 可以再用矩阵的形式表示成如下:

定理 8.2.3 任意方阵 A 相似于有理典范型中唯一的一个矩阵.

推论 8.2.4 同一域 F 上的两个矩阵相似,当且仅当它们有相同的初等因子.

我们不想详细地讨论如何求得矩阵的有理典范型,因为我们对有理典范型研究的主要目的在于把它当作一个理论工具. 但是为了对有理典范型有一个具体的了解,下面给出一个有理典范型的例子.

例 8.2.1 τ 是 R^7 上的一线性算子,设 τ 有极小多项式 $m_\tau(x)=(x-1)(x^2+1)^2$.

解 $x-1$ 和$(x^2+1)^2$ 是初等因子,对于初等因子列,我们有以下两种可能性:

1) $x-1,(x^2+1)^2,x^2+1$;

2) $x-1,x-1,x-1,(x^2+1)^2$.

相应地,它们有以下两种有理典范型.

$$1)\begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix},\quad 2)\begin{pmatrix} -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & -2 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

8.3　算子的本征值与本征向量

8.3.1　算子的本征多项式

引理 8.3.1　如果方阵 M 有分块对角型 $M=\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$，其中 A 和 B 是方阵，那么 $\det(M)=\det(A)\det(B)$.

令 $U(p(x))$ 是多项式 $p(x)=a_0+a_1x+\cdots+a_{n-1}x^{n-1}+x^n$ 的友矩阵，且令

$$A=xI-U(p(x))=\begin{pmatrix} x & 0 & \cdots & 0 & a_0 \\ -1 & x & \cdots & 0 & a_1 \\ 0 & -1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & x & a_{n-2} \\ 0 & 0 & \cdots & -1 & x+a_{n-1} \end{pmatrix}.$$

我们把 A 记作 $A=A(x;a_0,\cdots,a_{n-1})$，然后求出矩阵的行列式与系数 a_i 的相关性，从而计算出这个矩阵的行列式，我们首先看一些简单的情况.

$$\det(A(x;a_0,a_1))=\begin{vmatrix} x & a_0 \\ -1 & x+a_1 \end{vmatrix}=x(x+a_1)+a_0=a_0+a_1x+x^2=p(x);$$

$$\det(A(x;a_0,a_1,a_2))=\begin{vmatrix} x & 0 & a_0 \\ -1 & x & a_1 \\ 0 & -1 & x+a_2 \end{vmatrix}$$

$$=x\begin{vmatrix} x & a_1 \\ -1 & x+a_2 \end{vmatrix}+a_0\begin{vmatrix} -1 & x \\ 0 & -1 \end{vmatrix}$$

$$=x[x(x+a_2)+a_1]+a_0=a_0+a_1x+a_2x^2+x^3=p(x).$$

通常，沿着第一行展开得

$$\det(A(x;\ a_0,\cdots,a_{n-1})) = x\det(A(x;\ a_1,\cdots,a_{n-1})) + (-1)^{n+1}(-1)^{n-1}a_0$$
$$= x\det(A(x;\ a_1,\ \cdots,\ a_{n-1})) + a_0.$$

因此,通过归纳论证,就有了以下

引理 8.3.2　如果 $U(p(x))$是多项式 $p(x)$的友矩阵,那么

$$\det(xI-U(p(x))) = p(x).$$

引理 8.3.1 与引理 8.3.2 结合起来就有了以下

定理 8.3.3　如果 R 是 $\tau\in L(V)$的有理典范型,那么

$$f_\tau(x) = \det D(xI-R) = \prod_{i,j} p_i^{e_{i,j}}(x).$$

定义 8.3.1　$f_\tau(x) = \det(xI-R) = \prod_{i,j} p_i^{e_{i,j}}(x)$称为算子 τ 的**本(特)征多项式**.

我们常常先把本征多项式定义在矩阵上,再定义在线性算子上. 方阵 A 的本征多项式定义为 $f_A(x) = \det(xI-A)$.

定理 8.3.4　1) 如果 A 与 B 相似,那么 $f_A(x) = f_B(x)$. 因此,在相似性之下,本征多项式是不变本征多项式.

2) 算子 τ 的特征多项式与任意表示 τ 的矩阵的特征多项式相同.

3) 算子 τ 的特征多项式是 τ 的初等因子的积.

虽然在相似性下,特征多项式是不变特征多项式(和极小多项式一样),但是有着相同特征多项式的矩阵

$$A = \begin{pmatrix} \delta & 0 \\ 0 & \delta \end{pmatrix} \text{与} B = \begin{pmatrix} \delta & 0 \\ 1 & \delta \end{pmatrix}$$

却不相似. 这说明特征多项式并不是完全不变式.

8.3.2　本征值与本征向量

定理 8.3.5　λ 是线性算子 $\tau\in L(V)$的本征多项式 $f_\tau(x)$的根$\Leftrightarrow\det(\lambda I-R) = 0$.

特别地,若 $\dim(V) = n$,那么 τ 的有理标准型 R 为 $n\times n$,所以 $\det(\lambda I-R) = 0$ 成立当且仅当存在非零向量 $\xi\in F^n$ 使得$(\lambda I-R)\xi = 0$,或 $\tau_R(\xi) = \lambda\xi$.

若 $v\in V$ 是使得$(v)_B = x$ 的非零向量,这里 B 是 R 用来表示 τ 的有序基,那么上式等价于 $\tau(v) = \lambda v$.

定义 8.3.2　设 $\tau\in L(V)$是一个线性算子,数 $\lambda\in F$ 是 τ 的一个**本征值**,如果存在非零向量 $v\in V$,使得 $\tau(v) = \lambda v$,那么 v 就是 τ 的属于本征值 λ 的**本征向量**.

若 A 是 F 上的矩阵,那么 $\lambda\in F$ 是 A 的一个**本征值**,如果存在非零列向量 x,使得 $A\xi=\lambda\xi$,那么 ξ 就是 A 的属于本征值 λ 的**本征向量**.

定义 8.3.3　所有属于本征值 λ 的本征向量所成的集合，加之零向量构成了 V 的一个子空间，称为 λ 的**本征子空间**.

我们将用 δ_λ 表示本征值 λ 的本征空间，它可应用于线性算子和矩阵.

例 8.3.1　设 A 是代数闭域 F(比如复数域 $\mathbf{C}$)上的 $n \times n$ 矩阵. 因此，一个本征多项式的所有根都在 F 中，则 $\det(A)$ 是 A 的本征值的积.

以下定理总结了本征值和本征向量的基本性质：

定理 8.3.6　1) 数 $\lambda \in F$ 是 $\tau \in L(V)$ 的本征值，当且仅当 λ 是 τ 的本征多项式 $C_\tau(x)$ 的根.

2) 数 $\lambda \in F$ 是 $\tau \in L(V)$ 的本征值，当且仅当 λ 是 τ 的极小多项式 $m_\tau(x)$ 的根.

3) 数 $\lambda \in F$ 是 $\tau \in L(V)$ 的本征值，当且仅当 λ 是表示 τ 的任意矩阵的本征值.

4) 在相似性下，矩阵的本征值是一个不变量.

5) 若 λ 是矩阵 A 的一个本征值，那么本征子空间 δ_λ 是齐次线性方程组 $(\lambda I - A)X = 0$ 的解空间.

证明　1) 已证.

2) 本征多项式 $C_\tau(x)$ 和极小多项式 $m_\tau(x)$ 的素因子相同.

3) 对于任意非零的 $v \in V$，λ 是 τ 的一个本征值，当且仅当 $\tau(v) = \lambda v$. 现在，设 $\dim(V) = d$，令 B 是 V 的一个有序基，且设 $\varphi_B: V \to F^d$ 是由 $\varphi_B(u) = (u)_B$ 定义的同构. 那么，如果 $A = (\tau)_B$，我们有 $\tau = (\varphi_B)^{-1}\tau_A\varphi_B$(见图 8-1). 所以 $\tau(v) = \lambda v$ 等价于 $(\varphi_B)^{-1}\tau_A\varphi_B(v) = \lambda v = (\varphi_B)^{-1}\lambda\varphi_B(v)$，或 $\tau_A\varphi_B(v) = \lambda(v)$. 这表明 λ 是 A 的一个本征值.

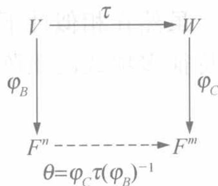


图 8-1

4) 可以从 3) 中得出，5) 是显然的.

例 8.3.2　若 $\tau, \sigma \in L(V)$，那么 $\tau\sigma$ 和 $\sigma\tau$ 的本征值相同，若 $\tau\sigma = \sigma\tau$，那么 τ 和 σ 有一个公共本征向量.

例 8.3.3　如果 λ 是 τ 的一个本征值，那么对于任意多项式 $p(x)$，$p(\lambda)$ 是 $p(\tau)$ 的本征值. 同样，如果 $\lambda \neq 0$，那么 λ^{-1} 是 τ^{-1} 的本征值.

8.4　幂零算子的标准分解

若当分解告诉我们，复数域上每一个 n 阶矩阵都可以表示成一个可对角化的矩阵与一个幂零矩阵的和. 可对角化部分的形式是简单的，我们先来讨论幂零矩阵的标准形式.

8.4.1　幂零若当块

设 σ 是向量空间 V 的一个幂零算子，那么存在一个正整数 r，使得 $\sigma^r = \theta$. 我们可以

进一步假定 r 是使 $\sigma^r=\theta$ 的最小正整数，那么 σ 的最小多项式是 x^r，于是存在一个向量 ξ_0，使得 $\sigma^r(\xi_0)=0$，而 $\sigma^{r-1}(\xi_0)\neq 0$，则向量组 $\xi_0,\sigma(\xi_0),\cdots,\sigma^{r-1}(\xi_0)$ 线性无关. 令 W 是由这一组向量所生成的子空间，那么 $\dim W=r$，而向量组作成 W 的一个基.

定义 8.4.1　设 σ 是向量空间 V 的一个线性算子，子空间 W 叫做关于 σ 的一个循环子空间，简称 **σ-循环子空间**，如果存在一个向量 ξ_0 和一个正整数 r，使得

1) $\xi_0,\sigma(\xi_0),\cdots,\sigma^{r-1}(\xi_0)$ 构成 W 的一个基；

2) $\sigma^r(\xi_0)=0$.

这时 ξ_0 叫做循环子空间 W 的一个**生成向量**，而 $\xi_0,\sigma(\xi_0),\cdots,\sigma^{r-1}(\xi_0)$ 叫做 W 的一个**循环基.**

显然，一个 σ-循环子空间 W 在 σ 之下不变，并且对于任意 $\xi\in W$，都有 $\sigma^r(\xi_0)=0$，这里 $r=\dim W$.

定义 8.4.2　令 W 是一个 σ-循环子空间，而 $\xi_0,\sigma(\xi_0),\cdots,\sigma^{r-1}(\xi_0)$ 是 W 的一个循环基，那么 σ 在 W 上的限制 $\sigma|_W$ 是 W 的一个幂零算子，并且关于这个基的矩阵是如下形状的一个 r 阶矩阵：

$$N_r=\begin{pmatrix}0&0&0&\cdots&0\\1&0&0&&\\0&1&0&\ddots&\vdots\\\vdots&&\ddots&\ddots&0\\0&\cdots&0&1&0\end{pmatrix}$$

矩阵 N_r 叫做一个 r 阶**幂零若当矩阵**或**幂零若当块.**

我们要证明，对于 n 维向量空间 V 的每一幂零线性算子，V 可以分解为一些 σ-循环子空间 $W_1,\cdots,W_n$ 的直和. 于是在每一个子空间 W_i 内选取一个循环基，凑起来成为 V 的一个基，σ 关于这个基的矩阵有形状

$$\begin{pmatrix}N_{r_1}&&&\\&N_{r_2}&&\\&&\ddots&\\&&&N_{r_s}\end{pmatrix}.$$

8.4.2　幂零线性算子

这里每一 N_{r_i} 是一个 r_i 阶幂零若当块，$r_1+\cdots+r_s=n$.

引理 8.4.1　设 σ 是向量空间 V 的一个幂零线性算子，而

$$h(x)=a_0+a_1x+\cdots+a_mx^m$$

是一个多项式,那么当且仅当 $a_0 \neq 0$ 时,线性变换 $h(\sigma)$ 有逆算子. 当 $h(\sigma)$ 可逆时,$h(\sigma)$ 的逆算子也是 σ 的一个多项式.

证明　首先注意,如果 σ 是一个幂零算子,$\sigma^r = 0$,那么 $\tau = a_0 + a_1\sigma + \cdots + a_m\sigma^m$ 也是一个幂零算子. 这只要计算一下等式右端的 r 次方,就可以看出 $\tau^r = \theta$. 这样,当 $a^0 = 0$ 时,$h(\sigma)$ 没有逆算子.

现在设 $a_0 \neq 0$,我们有 $h(\sigma) = a_0 + \tau$,这里 $\tau = a_0 + a_1\sigma + \cdots + a_m\sigma^m$. 由上面的说明得,$\tau$ 是幂零变换. 设 $\tau^r = 0$,直接计算可知

$$a_0 + \tau\left(\frac{1}{a_0} - \frac{1}{a_0^2}\tau + \cdots + (-1)^{r-1}\frac{1}{a_0^r}\tau^{r-1}\right) = \iota,$$

这里 ι 是 V 的单位算子,所以 $h(\sigma)$ 可逆. 因为 τ 是 σ 的多项式,所以

$$h(\sigma)^{-1} = \frac{1}{a_0} - \frac{1}{a_0^2}\tau + \cdots + (-1)^{r-1}\frac{1}{a_0^r}\tau^{r-1},$$

也是 σ 的一个多项式.

引理 8.4.2　设 σ 是向量空间 V 的一个幂零线性算子. W 是一个 r 维 σ-循环子空间,$\xi \in W$. 如果存在一个整数 k,$0 \leqslant k \leqslant r$,使得 $\sigma^{r-k}(\xi) = 0$,那么存在 $\eta \in W$,使得 $\xi = \sigma^k(\eta)$.

证明　令 $\xi_0, \sigma(\xi_0), \cdots, \sigma^{r-1}(\xi_0)$ 是 W 的一个循环基,于是

$$\xi = a_0\xi_0 + a_1\sigma(\xi_0), \cdots, + a_{r-1}\sigma^{r-1}(\xi_0).$$

因为 $\sigma^{r-k}(\xi) = 0$,所以 $a_0\sigma^{r-k}(\xi_0) + \cdots + a_{k-1}\sigma^{r-1}(\xi_0) = 0$. 然而 $\sigma^{r-k}(\xi_0), \cdots, \sigma^{r-1}(\xi_0)$ 线性无关,所以 $a_0 = \cdots = a_{k-1} = 0$,从而

$$\xi = a_k\sigma^k(\xi_0) + \cdots + a_{r-1}\sigma^{r-1}(\xi_0) = \sigma^k(a_k(\xi_0) + \cdots + a_{r-1}\sigma^{r-k-1}(\xi_0)),$$

取 $\eta = a_k(\xi_0) + \cdots + a_{r-1}\sigma^{r-k-1}(\xi_0)$,那么就有 $\xi = \sigma^k(\eta)$.

引理 8.4.3　设 σ 是 n 维向量空间 V 的一个幂零线性算子,x^r 是 σ 的最小多项式,令 W 是一个 r 维 σ-循环子空间,那么存在 W 的一个余子空间 W': $V = W \oplus W'$,并且 W' 也在 σ 之下不变.

证明　取 W 是 V 中具有下列两个性质的一个极大子空间:

1) $W \cap W' = \{0\}$;

2) W' 在 σ 之下不变.

具有性质 1),2)的子空间显然是存在的,因为零空间就是一个. 因此,这样的极大子空间是存在的. 我们证明 $V = W + W'$,从而由性质 1)知,这个和是直和.

如果 $V \neq W + W'$,那么存在 $\xi \in V$,而 $\xi \notin W + W'$,因为 $\sigma(\xi) = 0$,所以存在一个整数 k,$0 < k \leqslant r$,使得 $\sigma^k(\xi) \in W + W'$,而对于任何小于 k 的正整数 i 来

说，$\sigma^i(\xi) \notin W + W'$.

我们有 $\sigma^k(\xi) = \xi + \xi', \xi \in W, \xi' \in W'$，于是

$$\sigma^{r-k}(\xi) + \sigma^{r-k}(\xi') = \sigma^r(\xi) = 0.$$

因为 W 和 W' 都在 σ 之下不变，所以 $\sigma^{r-k}(\xi) \in W, \sigma^{r-k}(\xi') \in W'$，从而

$$\sigma^{r-k}(\xi) = -\sigma^{r-k}(\xi') \in W \cap W'.$$

由此得出 $\sigma^{r-k}(\xi) = 0$. 于是由引理 8.4.2 得，存在向量 $\eta \in W$，使得 $\xi = \sigma^k(\eta)$. 因此 $\sigma^k(\xi) = \xi + \xi', \xi \in W, \xi' \in W'$，可以写成 $\sigma^k(\xi) = \sigma^k(\eta) + \xi$，令 $\zeta = \xi - \eta$，那么

$$\sigma^k(\zeta) = \sigma^k(\xi) - \sigma^k(\eta) = \xi',$$

因为 W' 在 σ 之下不变，所以对于任意整数 $m \geqslant k$ 来说，都有 $\sigma^m(\zeta) \in W'$.

另一方面，如果 $i < k$，那么 $\sigma^i(\xi) \notin W + W'$，然而 $\sigma^i(\eta) \in W$，所以

$$\sigma^i(\zeta) = \sigma^i(\xi) - \sigma^i(\eta) \notin W + W',$$

特别由此得出 $\zeta \notin W'$.

令 U 是由 $\zeta, \sigma(\zeta), \cdots, \sigma^{k-1}(\zeta)$ 所生成的子空间，令 $W_1 = U + W'$，因为 $\sigma^k(\zeta) \in W'$，所以 W_1 在 σ 之下不变. 又因为 $\zeta \notin W'$，所以 $\dim W_1 > \dim W'$，由 W' 的极大性得出 $W \cap W_1 \neq \{0\}$.

所以存在 $\alpha \in W \cap W_1, \alpha \neq 0$. 于是 α 可以写成

$$\alpha = b_0(\zeta) + b_1\sigma(\zeta) + \cdots + b_{k-1}\sigma^{k-1}(\zeta) + \alpha_2, \alpha_2 \in W',$$

$b_0, b_1, \cdots, b_{k-1}$ 不能全为零，否则 $\alpha = \alpha_2$，从而 $\alpha \in W \cap W' = \{0\}$. 这与 α 的取法矛盾. 令 b_i 是第一个不等于零的系数，那么

$$\alpha = b_i\sigma^i(\zeta) + \cdots + b_{k-1}\sigma^{k-1}(\zeta) + \alpha_2 = (b_i + \cdots + b_{k-1}\sigma^{k-i-1})\sigma^i(\zeta) + \alpha_2,$$

由引理 8.4.1 知，线性算子 $(b_i + \cdots + b_{k-1}\sigma^{k-i-1})$ 可逆，并且它的逆算子 τ 也是 σ 的多项式，因此 W 和 W' 都在 τ 之下不变. 对上面等式两端施行线性变换，注意到 $\alpha \in W$，我们有

$$\sigma^i(\zeta) + \tau(\alpha_2) = \tau(\alpha) \in W,$$

又 $\tau(\alpha_2) \in W'$，所以 $\sigma^i(\zeta) = \tau(\alpha) - \tau(\alpha_2) \in W + W'$，这与若 $i < k, \sigma^i(\zeta) = \sigma^i(\xi) - \sigma^i(\eta) \notin W + W'$ 矛盾，这就证明了 $V = W + W'$.

8.4.3 向量空间的分解

定理 8.4.4 设 σ 是 n 维向量空间 V 的一个幂零线性算子，那么 V 可以分解为 σ-循环子空间的直和：

$$V=W_1\oplus\cdots\oplus W_s.$$

令 $r_i=\dim W_i, i=1,2,\cdots,s$,我们有 $r_1\geqslant r_2\geqslant\cdots\geqslant r_s$.

证明　对 V 的维数 n 作数学归纳法. 当 $n=0$ 时,定理显然成立.

假设 $n>0$,并且对于维数小于 n 的向量空间来说,定理成立.

设 $\dim V=n$,令 x^r 是 σ 的最小多项式,那么存在 $\xi\in V$,使得 $\sigma^{r_1}(\xi)=0$,而 $\sigma^{r_1-1}(\xi)\neq 0$. 令 W 是由 ξ 所生成的 σ-循环子空间. 由引理 8.4.3 得,存在一个在 σ 之下不变的子空间 W,使得 $V=W_1\oplus W$. σ 在 W 上的限制 $\sigma|_W$ 是 W 的一个幂零算子,而 $\dim W<n$. 由归纳法的假设,W 可以分解为 $\sigma|_W$-循环子空间的直和

$$W=W_2\oplus\cdots\oplus W_s,$$

并且它们的维数满足关系 $r_2\geqslant\cdots\geqslant r_s, r_i=\dim W_i, i=2,3,\cdots,s$.

每一子空间 W_i 自然也是 σ-循环子空间,于是

$$V=W_1\oplus\cdots\oplus W_s.$$

因为 r_1 是 σ 的最小多项式的次数,而 r_2 是 $\sigma|_W$ 的最小多项式的次数,所以 $r_1\geqslant r_2$,在定理 8.4.4 里,对于每一 σ-循环子空间 W_i,取一个循环基

$$\xi_i, \sigma^k(\xi_i), \cdots, \sigma^{r_1-1}(\xi_i), i=1,2,\cdots,s$$

凑起来成为 V 的基,那么 σ 关于这个基的矩阵有形状

$$\begin{pmatrix} N_{r_1} & & & \\ & N_{r_2} & & \\ & & \ddots & \\ & & & N_{r_s} \end{pmatrix},$$

这里 N_{r_i} 是一个 r_i 阶幂零若当块,$r_1\geqslant\cdots\geqslant r_s$. 于是我们得到

定理 8.4.5　每一 n 阶幂零矩阵都与一个形如

$$N=\begin{pmatrix} N_{r_1} & & & \\ & N_{r_2} & & \\ & & \ddots & \\ & & & N_{r_s} \end{pmatrix}$$

的矩阵相似,这里 N_{r_i} 是一个 r_i 阶幂零若当块,$r_1\geqslant\cdots\geqslant r_s$.

下边考虑,出现在定理 8.4.4 的空间分解中的循环子空间的维数序列 $r_1\geqslant\cdots\geqslant r_s$ 是不是唯一确定的. 我们将给出肯定的回答.

首先注意到,设 W 是一个 σ-循环子空间,$\dim W=r$. 令 s 是一个整数,$0<s\leqslant r$,那

么子空间 $\sigma^i(W)$ 的维数是 $r-s$.

事实上，令 $\xi,\sigma(\xi),\cdots,\sigma^{r-1}(\xi)$ 是 W 的一个循环基，那么

$$\sigma^s(\xi),\sigma^{s+1}(\xi),\cdots,\sigma^{r-1}(\xi)$$

就是 $\sigma^s(W)$ 的一个基.

8.4.4　分解的唯一性

定理 8.4.6　设 σ 是 n 维向量空间 V 的一个幂零线性算子. 如果有两种方式将 V 分解为 σ-循环子空间的直和：

$$V = W_1 \oplus \cdots \oplus W_s = U_1 \oplus \cdots \oplus U_s,$$

并且 $\dim W_i = r_i, i = 1,2,\cdots,s, \dim U_j = k_j, j = 1,2,\cdots,t$，满足条件 $r_1 \geqslant \cdots \geqslant r_s$，$k_1 \geqslant \cdots \geqslant k_t$，那么 $s = t, r_i = k_i, i = 1,2,\cdots,s$.

证明　如果命题不成立，令 i 是第一个使 $r_i \neq k_i$ 的整数，不妨设 $r_i > k_i$，考虑 $\sigma^{k_i}(V)$. 因为每一子空间 W_j 都在 σ 之下不变，所以

$$\sigma^{k_i}(V) = \sigma^{k_i}(W_1) \oplus \cdots \oplus \sigma^{k_i}(W_s).$$

注意到，如果 $r_j > k_i$，那么 $\dim\sigma^{k_i}(W_j) = r_j - k_i$. 因此有

$$\dim\sigma^{k_i}(V) \geqslant (r_1 - k_i) + \cdots + (r_i - k_i).$$

另一方面，每一子空间 U_j 也在 σ 之下不变，并且当 $j \geqslant i$ 时，$\sigma^{k_i}(U_j) = 0$，则

$$\sigma^{k_i}(V) = \sigma^{k_i}(U_1) \oplus \cdots \oplus \sigma^{k_i}(U_{i-1}),$$

从而 $\dim\sigma^{k_i}(V) = (k_1 - k_i) + \cdots + (r_{i-1} - k_i)$. 但 $r_1 = k_1, \cdots, r_{i-1} = k_{i-1}$，所以有

$$\dim\sigma^{k_i}(V) \geqslant (r_1 - k_i) + \cdots + (r_{i-1} - k_i).$$

比较上式，又 $r_i - k_i > 0$，这就导致矛盾.

设 σ 是 n 维向量空间 V 的一个幂零线性算子. 我们把出现在 V 关于 σ 的循环子空间分解式中的唯一确定的一组正整数 $r_1 \geqslant \cdots \geqslant r_s$ 叫做 σ 的不变指数. 同样地，如果 A 是一个 n 阶幂零矩阵，由定理 8.4.5 知，A 与一个形如

$$N = \begin{pmatrix} N_{r_1} & & & \\ & N_{r_2} & & \\ & & \ddots & \\ & & & N_{r_s} \end{pmatrix}$$

的矩阵相似. 出现在矩阵 N 里的幂零若当块的阶 $r_1 \geqslant \cdots \geqslant r_s$ 是由 A 唯一确定的. 正整数序列 $r_1 \geqslant \cdots \geqslant r_s$ 也叫做矩阵 A 的不变指数.

8.5　算子的若当标准型

8.5.1　凯莱—哈密顿定理

定理 8.5.1　设 $\lambda_1, \cdots, \lambda_k$ 是线性算子 $\tau \in L(V)$ 的互异的本征值，则 $\delta_{\lambda_i} \cap \delta_{\lambda_j} = \{0\}$. 而且，属于互异本征值的本征向量是线性无关的，即，若对于 $i = 1, \cdots, k, v_i \in \delta_{\lambda_i}$，那么向量 $\{v_1, \cdots, v_k\}$ 是线性无关的.

证明　令对于 $i = 1, \cdots, k, v_i \in \delta_{\lambda_i}$，这里 $\lambda_1, \cdots, \lambda_k$ 是 τ 的互异的本征值. 我们想要说明 v_i 是线性无关的，假设 v_i 是线性相关的，我们也可以设（必要时重新编号）在这些向量所成的等于零的非平凡线性组合中，方程

$$r_1 v_1 + \cdots + r_j v_j = 0$$

是最短的这种方程（即所含的项最少）. 代入 τ 得

$$r_1 \tau(v_1) + \cdots + r_j \tau(v_j) = 0 \text{ 或 } r_1 \lambda_1 v_1 + \cdots + r_j \lambda_j v_j = 0.$$

这里 $r_1 v_1 + \cdots + r_j v_j = 0$，乘以 λ_1，并从上式中减去 $r_1 v_1 + \cdots + r_j v_j = 0$，得

$$r_2(\lambda_2 - \lambda_1) v_2 + \cdots + r_j(\lambda_j - \lambda_1) v_j = 0.$$

上式比 $r_1 v_1 + \cdots + r_j v_j = 0$ 所含的项少，所以其系数必都为 0，又因为 λ_i 是互异的本征值，所以可以推出对于 $i = 2, \cdots, j, r_i = 0$，因此 r_1 也等于 0. 这一矛盾说明 v_i 是线性无关的.

计算线性算子 τ 的本征值的方法之一是首先用矩阵 A 表示 τ，然后解特征方程 $C_A(x)=0$，但当 $C_A(x) = \dim(V) \geqslant 3$ 时，我们很可能解不出这个多项式方程，从而逼近矩阵的本征值的技巧就成了应用线性代数研究的一个重要领域.

因为线性算子 τ 的特征多项式 $C_\tau(x)$ 是其初等因子的积，又因为 τ 的极小多项式是积

$$m_\tau(x) = p_1^{e_1}(x) \cdots p_n^{e_n}(x),$$

这样，我们推出 $m_\tau(x) \mid C_\tau(x)$. 这一重要结论称为凯莱—哈密顿定理.

定理 8.5.2　设 $\tau \in L(V)$.

1）极小多项式 $m_\tau(x)$ 与特征多项式 $C_\tau(x)$ 有相同的素因子.

2）（凯莱-哈密顿定理）　$m_\tau(x) \mid C_\tau(x)$，或等价地，τ 满足其自身的特征多项式.

8.5.2　若当标准型

有理标准型的优点之一是有限维向量空间上任意线性算子 τ 都有一个有理标准型，即，有理标准型中所有矩阵所成的集合构成了一个有理标准型集（至少对于对角线

上分块的阶来说). 但矩阵的有理标准型远远没有达到我们所想像的简单标准型集那样简单.

当然,在某些重要情形中,我们还有比有理标准型更好的例子.

特别地,考虑线性算子 $\tau \in L(V)$,其极小多项式因子可分解为线性因子的积

$$m_\tau(x) = (x-\lambda_1)^{e_1} \cdots (x-\lambda_n)^{e_n}.$$

定义 8.5.1　当域 F 上一个多项式因子分解为线性算子的一个积时,我们就说这个多项式在 F 上**分裂**.

定义 8.5.2　域 F 上每个非零次多项式在 F 中都有一个根, F 称为是**代数闭域**.

因此,代数闭域上唯一的不可约多项式就是线性多项式,所以 F 上任何非常数多项式都在 F 上分裂.

例 8.5.1　复数 $\mathbf{C}$ 构成一个代数闭域,所以复向量空间上任意线性算子都含有在 $\mathbf{C}$ 上分裂的极小多项式.

从某种意义上说,有理标准型的"弱点"在于循环子模 $<v_{i,j}>$ 的基的选取,这里, $<vi, j>$ 的首一阶是初等因子 $p_i^{e_{i,j}}(x)$,通常,我们对它了解得很少. 因为 $<v_{i,j}>$ 是 V 的 τ 循环子空间,所以我们选取有序基

$$B_{i,j} = \{v_{i,j}, \tau_{i,j}(v_{i,j}), \cdots, {\tau_{i,j}}^{d_{i,j}-1}(v_{i,j})\},$$

但是,当极小多项式有形式 $m_\tau(x) = (x-\lambda_1)^{e_1} \cdots (x-\lambda_n)^{e_n}$,那么初等因子就有形式

$$p_i^{e_{i,j}}(x) = (x-\lambda_i)^{e_{i,j}}.$$

这样,我们可以更好地选取有序基. 观察到 $\dim(<v_{i,j}>) = \deg p_i^{e_{i,j}}(x)$,所以不难看出,集合

$$X_{i,j} = \{v_{i,j}, (\tau_{i,j}-\lambda_i)(v_{i,j}), \cdots, (\tau_{i,j}-\lambda_i)^{e_{i,j}-1}(v_{i,j})\}$$

是 $<v_{i,j}>$ 的有序基. 而且,在 $X_{i,j}$ 中,用 b_k 表示第 k 个基向量,对于 $k=0, \cdots, e_{i,j}-2$,

$$\begin{aligned}\tau_{i,j}(b_k) &= \tau_{i,j}[(\tau_{i,j}-\lambda_i)^k(v_{i,j})] = (\tau_{i,j}-\lambda_i+\lambda_i)[(\tau_{i,j}-\lambda_i)^k(v_{i,j})] \\ &= (\tau_{i,j}-\lambda_i)^{k+1}(v_{i,j}) + \lambda_i(\tau_{i,j}-\lambda_i)^k(v_{i,j}) = b_{k+1}+\lambda_i b_k.\end{aligned}$$

对于 $k = e_{i,j}-1$,运用 $(\tau_{i,j}-\lambda_i)^{k+1}(v_{i,j}) = (\tau_{i,j}-\lambda_i)^{e_{i,j}}(v_{i,j}) = 0$ 的类似运算,得 $\tau_{i,j}(b_{e_{i,j}-1}) = b_{e_{i,j}-1}$. 因此, $\tau_{i,j} = \tau|_{<v_{i,j}>}$ 的矩阵是 $e_{i,j} \times e_{i,j}$ 矩阵

$$g(\lambda_i, e_{i,j}) = \begin{bmatrix} \lambda_i & 0 & \cdots & 0 \\ 1 & \lambda_i & & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 1 & \lambda_i \end{bmatrix}.$$

定义 8.5.3　这个矩阵称为属于数 λ_i 的**若当块**.

定理 8.5.3 令算子 $\tau\in L(V)$ 的极小多项式在基域 F 上分裂,即

$$m_\tau(x)=(x-\lambda_1)^{e_1}\cdots(x-\lambda_n)^{e_n},$$

那么我们可以记作

$$V=(<v_{1,1}>\oplus\cdots\oplus<v_{1,k_1}>)\oplus\cdots\oplus(<v_{n,1}>\oplus\cdots\oplus<v_{n,k_n}>),$$

其中 $<v_{i,j}>$ 是 V 的 τ 循环子空间. $\tau_{i,j}=\tau|_{<v_{i,j}>}$ 的极小多项式是 V 的初等因子

$$\min(\tau_{i,j})=(x-\lambda_i)^{e_{i,j}},$$

其中 $e_i=e_{i,1}\geqslant e_{i,2}\geqslant\cdots\geqslant e_{i,k_i}$. 这些初等因子是由 τ 唯一确定的,而且集合

$$X_{i,j}=\{v_{i,j},(\tau_{i,j}-\lambda_i)(v_{i,j}),\cdots,(\tau_{i,j}-\lambda_i)^{e_{i,j}-1}(v_{i,j})\}$$

是 $<v_{i,j}>$ 的有序基. τ 的关于有序基 $X=\{X_{1,1},\cdots,X_{n,k_n}\}$ 的矩阵是分块矩阵

$$(\tau)_X=\begin{pmatrix} g(\lambda_1,e_{1,1}) & & & & & & \\ & \ddots & & & & & \\ & & g(\lambda_1,e_{1,k_1}) & & & & \\ & & & \ddots & & & \\ & & & & \ddots & & \\ & & & & & g(\lambda_n,e_{n,k_n}) & \\ & & & & & & \ddots \\ & & & & & & & g(\lambda_n,e_{n,k_n}) \end{pmatrix}_{\text{block}}$$

等式右边的矩阵称为 τ 的**若当标准型**.

对于代数闭域 F,若当标准型中的矩阵集确是相似性的一个标准型集(至少对于若当分块的阶来说). 换句话说,F 上每个矩阵恰与若当标准型中的一个矩阵相似(同样对于若当分块的阶而言).

如果 τ 有若当标准型 g,那么 g 的对角元正好是特征多项式 $C_\tau(x)$ 的根,而且 g 中每个对角元出现的次数就是作为特征多项式的根的元素的重数.

8.5.3 算子的对角化

定义 8.5.4 若 $\lambda\in F$ 是线性算子 τ 的本征值,那么 λ 的重数,即特征多项式 $C_\tau(x)$ 的根,称为 λ 的**代数重数**.

本征空间 δ_λ 的维数称为 λ 的**几何重数**.

定理 8.5.4 $\tau\in L(V)$ 的本征值 λ 的几何重数小于等于其代数重数.

证明 设 λ 是 τ 的一个本征值,则 $m_\tau(x)=(x-\lambda)^e p(x)$,其中 $x-\lambda$ 不整除 $p(x)$.

考虑 τ 的有理标准型. 在 V 的准素分解中,有 $V = V_{p_1} \oplus \cdots \oplus V_{p_n}$. 设 $V_{p_1} = \{v \in V \mid (x-\lambda)^e v = 0\}$, 这个准素子模的循环分解是 $V_{p_1} = \langle v_1 \rangle \oplus \cdots \oplus \langle v_k \rangle$,且有初等因子 $\min(\tau_j) = (x-\lambda)^{e_j}$,其中 $\tau_j = \tau|_{\langle v_j \rangle}$,且 $e = e_1 \geqslant e_2 \geqslant \cdots \geqslant e_k$. λ 的代数重数是

$$\text{alg. mult.} = \sum_{j=1}^{k} e_j = \sum_{j=1}^{k} \dim(\langle v_j \rangle) = \dim(V_{p_1}),$$

则 $$v \in \delta_\lambda \Rightarrow \tau(v) = \lambda v \Rightarrow (\tau - \lambda)(v) = 0 \Rightarrow (\tau - \lambda)^e(v) = 0 \Rightarrow v \in V_{p_1}.$$

因此 $\delta_\lambda \subset V_{p_1}$,即 τ 的属于 λ 的本征向量都在 V_{p_1} 中. 对于 $\forall j = 1, \cdots, k$,集合

$$X_j = \{v_j, (\tau_j - \lambda)(v_j), \cdots, (\tau_j - \lambda)^{e_j - 1}(v_j)\}$$

是 $\langle v_j \rangle$ 的基,且 $(\tau_j - \lambda)^{e_j}(v_j) = 0$. 在 V_{p_1} 中,若 $u = \sum_{i=0}^{e_j-1} r_i (\tau - \lambda)^i (v_j)$ 是一本征向量,则 $(\tau - \lambda)(u) = 0$,即

$$0 = \sum_{i=0}^{e_j-1} r_i (\tau - \lambda)^{i+1}(v_j) = \sum_{i=0}^{e_j-2} r_i (\tau - \lambda)^{i+1}(v_j).$$

所以对于 $i = 0, \cdots, e_j - 2$, $r_i = 0$,有 $u = r_{e_j-1}(\tau - \lambda)^{e_j-1}(v_j)$.

因此,$\langle v_j \rangle$中唯一的本征向量是向量$(\tau - \lambda)^{e_j-1}(v_j)$的数量倍. 这说明 λ 的几何重数是构成 V_{p_1} 的 τ 循环子空间$\langle v_j \rangle$的数 k,因为代数重数是这些 τ 循环子空间的维数的和.

8.5.4　可对角化算子的特征

现在给出可对角化线性算子的几个不同特征,即可以由对角阵表示的算子.

定理 8.5.5　算子 $\tau \in L(V)$是可对角化的,当且仅当存在 V 的一个基,且这个基完全由 τ 的本征向量组成,即,当且仅当 $\dim(\delta_{\lambda_1} \oplus \cdots \oplus \delta_{\lambda_k}) = \dim(V)$,其中 $\lambda_1, \cdots, \lambda_k$ 是 τ 的互异的本征值.

若当标准型给出可对角化算子的另一特征. 设 τ 可对角化,

$$(\tau)_{B'} = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_k \end{pmatrix}.$$

这里,对角元不一定要是互异的. 设 $\lambda_{i_1}, \cdots, \lambda_{i_s}$ 是互异的对角元,那么 τ 的极小多项式是 $m_\tau(x) = \prod_j (x - \lambda_{i_j})$,也是$(\tau)_B$ 的极小多项式. 因此,$m_\tau(x)$ 是互异的线性因子的积.

反之,设 $\tau\in L(V)$有性质:$m_\tau(x)$是互异的线性因子的积,令 $m_\tau(x)=(x-\lambda_1)\cdots(x-\lambda_s)$,这里 λ_i 是互异的,那么 τ 有若当标准型,而且所有初等因子都有形式 $x-\lambda_i$,所以 $\min(\tau_{i,j})=x-\lambda_i$.

换言之,$<v_{i,j}>$中所有向量都是本征向量,因此 τ 是可对角化的,这就有了以下

定理 8.5.6　有限维向量空间上线性算子 $\tau\in L(V)$是可对角化的,当且仅当 τ 的极小多项式是线性因子的积.

8.6　射影代数

8.6.1　射影

为了求可对角化算子的另一特征,我们转而讨论一种特殊的算子.

定义 8.6.1　设 $V=S+S^c$. 由 $\rho(s+s^c)=s$ 定义的映射 $\rho: V\to V$ 是 V 上的一个线性算子,称为 **S 上沿 S^c 的射影**,这里 $s\in S, s^c\in S^c$.

射影的概念如图 8-2 所示.

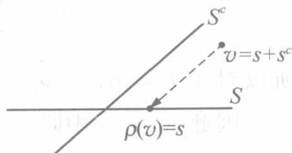


图 8-2

定理 8.6.1　1) 设 ρ 是 S 上沿 S^c 的射影,则 $\mathrm{Im}(\rho)=S$, $\mathrm{Ker}(\rho)=S^c$;$V=\mathrm{Im}(\rho)\oplus\mathrm{Ker}(\rho)$;$v\in\mathrm{Im}(\rho)\Leftrightarrow\rho(v)=s$.

2) 反之,若 $\sigma\in L(V)$ 有性质 $V=\mathrm{Im}(\rho)\oplus\mathrm{Ker}(\rho)$, $\sigma|_{\mathrm{Im}(\sigma)}=id$, 那么 σ 是 $\mathrm{Im}(\sigma)$ 上沿 $\mathrm{Ker}(\sigma)$ 的射影.

例 8.6.1　设 ρ 是射影,则 $\mathrm{Ker}(\iota-\rho)=\mathrm{Im}(\rho)$,$\mathrm{Im}(\iota-\rho)=\mathrm{Ker}(\rho)$.

射影算子(或简称为射影)在线性算子谱论中起着重要作用,我们将在后面进行讨论.现在我们转而讨论这些算子的一些基本性质.

定理 8.6.2　线性算子 $\rho\in L(V)$是射影,当且仅当 ρ 是幂等算子,即 $\rho^2=\rho$.

证明　欲说明射影算子 ρ 在 S 上沿 S^c 是幂等算子,注意到对 $\forall s\in S$ 和 $s^c\in S^c$,

$$\rho^2(s+s^c)=\rho(\rho(s+s^c))=\rho(s)=s=\rho(s+s^c),$$

所以 $\rho^2=\rho$. 反之,如果 ρ 是幂等算子,设 $S=\{v\in V \mid \rho(v)=v\}$ 是一切由 ρ 固定的向量所成的集合,那么 $S\subset\mathrm{Im}(\rho)$.

同样,如果 $v\in\mathrm{Im}(\rho)$,那么对 $\forall w\in V, v=\rho(w)$,所以 $\rho(v)=\rho^2(w)=\rho(w)=v$. 因此,$\mathrm{Im}(\rho)\subset S$,所以 $S=\mathrm{Im}(\rho)$.

换句话说,$\rho|_{\mathrm{Im}(\rho)}=id$. 如果 $v\in\mathrm{Im}(\rho)\cap\mathrm{Ker}(\rho)$,我们有 $\rho(v)=v$ 和 $\rho(v)=0$,所以 $v=0$. 因此,$\mathrm{Im}(\rho)\cap\mathrm{Ker}(\rho)=\{0\}$.

最后,注意到对于 $v\in V, v=\rho(v)+(v-\rho(v))\in\mathrm{Im}(\rho)\cap\mathrm{Ker}(\rho)$,所以 $V=\mathrm{Im}(\rho)\oplus\mathrm{Ker}(\rho)$. 由定理 8.6.1 得证.

例 8.6.2　算子 $\tau\in L(V)$是幂零算子,如果对 $\forall n\in\mathbf{N}$,$\tau^n=0$.

例 8.6.3　τ 是幂零算子，那么 0 是 τ 的一个本征值，且是 τ 的唯一本征值.

8.6.2　射影代数

若 ρ 是射影，则 $\iota-\rho$ 也是射影，其中 ι 是 V 上的单位算子，因为

$$(\iota-\rho)^2=\iota^2-\iota\rho-\iota\rho+\rho^2=\iota-\rho.$$

不难看出，$\mathrm{Ker}(\iota-\rho)=\mathrm{Im}(\rho)$，$\mathrm{Im}(\iota-\rho)=\mathrm{Ker}(\rho)$. 因此，如果 ρ 是 S 上沿 S^c 的射影，那么 $\iota-\rho$ 也是 S^c 上沿 S 的射影.

定义 8.6.2　如果 $\rho\sigma=\sigma\rho=0$，那么两射影 $\rho,\ \sigma\in L(V)$ 是**正交的**，记为 $\rho\perp\sigma$.

$$\rho\perp\sigma\Leftrightarrow\mathrm{Im}(\rho)\subset\mathrm{Ker}(\sigma),\ \mathrm{Im}(\sigma)\subset\mathrm{Ker}(\rho).$$

以下例子说明在正交性定义中，有条件 $\rho\sigma=0$ 是不够的，因为可能会 $\rho\sigma=0$，而 $\sigma\rho$ 连射影都不是.

例 8.6.4　令 $V=F^2$，设 $D=\{(x,\ x)\mid x\in F\}$，$X=\{(x,\ 0)\mid x\in F\}$，$Y=\{(0,\ y)\mid y\in F\}$.

因此，在 F^2 中，D 是一条对角线，X 是 x 轴，Y 是 y 轴. 我们用符号 $\rho_{A,B}$ 表示 A 上沿 B 的射影，所以 $\rho_{D,x}\rho_{D,y}=\rho_{D,y}\neq\rho_{D,x}=\rho_{D,y}\rho_{D,x}$. 从而，如果 ρ 和 σ 是射影，那么积 $\rho\sigma$ 和 $\sigma\rho$ 都是射影，但是 $\rho\sigma\neq\sigma\rho$.

易见 $\rho_{y,x}\rho_{x,D}=0$（是一个射影），但是 $\rho_{x,D}\rho_{y,x}$ 不是射影，因此可能出现 $\rho\sigma$ 是射影（甚至是零射影），但 $\sigma\rho$ 不是射影.

如果 ρ 和 σ 是射影，并不能由此得出 $\rho+\sigma$，$\rho-\sigma$ 或 $\rho\sigma$ 是射影；$\rho+\sigma$ 是射影当且仅当 $(\rho+\sigma)^2=\rho+\sigma$，或

$$\rho\sigma+\sigma\rho=0.$$

上式左乘 ρ，得 $\rho\sigma+\rho\sigma\rho=0$，右乘 ρ，得 $\rho\sigma\rho+\sigma\rho=0$. 因此 $\rho\sigma=\sigma\rho$.

与 $\rho\sigma+\sigma\rho=0$ 一起，得 $2\rho\sigma=0$. 因此，若 $\mathrm{char}(F)\neq 2$，那么 $\rho\sigma=0$，$\sigma\rho=0$. 这说明若 $\rho+\sigma$ 是射影，那么 $\rho\perp\sigma$. 反之，若 $\rho\sigma=\sigma\rho=0$，那么 $(\rho+\sigma)^2=\rho+\sigma$，所以 $\rho+\sigma$ 是一个射影.

现在设 $\rho\sigma=\sigma\rho=0$，因此 $\rho+\sigma$ 是一个射影. 为了确定 $\rho+\sigma$ 的核，我们有

$$(\rho+\sigma)(v)=0\Rightarrow\rho(v)+\sigma(v)=0\Rightarrow\rho^2(v)+\rho\sigma(v)=0\Rightarrow\rho(v)=0.$$

所以 $\mathrm{Ker}(\rho+\sigma)\subset\mathrm{Ker}(\rho)$. 同理，$\mathrm{Ker}(\rho+\sigma)\subset\mathrm{Ker}(\sigma)$，因此

$$\mathrm{Ker}(\rho+\sigma)\subset\mathrm{Ker}(\rho)\cap\mathrm{Ker}(\sigma).$$

反包含是显然的，所以

$$\mathrm{Ker}(\rho+\sigma)=\mathrm{Ker}(\rho)\cap\mathrm{Ker}(\sigma).$$

至于 $\rho+\sigma$ 的象，已知 $\mathrm{Im}(\rho+\sigma)$ 是 V 中由射影 $\rho+\sigma$ 固定的一个向量集. 因此，

$$v \in \mathrm{Im}(\rho+\sigma) \Rightarrow v = (\rho+\sigma)(v) = \rho(v)+\sigma(v) \in \mathrm{Im}(\rho)+\mathrm{Im}(\sigma).$$

所以 $\mathrm{Im}(\rho+\sigma) \subset \mathrm{Im}(\rho)+\mathrm{Im}(\sigma)$. 但 $x \in \mathrm{Im}(\rho) \cap \mathrm{Im}(\sigma) \Rightarrow \rho(x) = x = \sigma(x) \Rightarrow x = \rho(x) = \rho^2(x) = \rho\sigma(x) = 0$,则 $\mathrm{Im}(\rho) \cap \mathrm{Im}(\sigma) = \{0\}$. 因此,$\mathrm{Im}(\rho+\sigma) \subset \mathrm{Im}(\rho) \oplus \mathrm{Im}(\sigma)$.

为了 确定反不等式,若 $v \in \mathrm{Im}(\rho) \oplus \mathrm{Im}(\sigma)$,那么对于 $r \in \mathrm{Im}(\rho)$,$s \in \mathrm{Im}(\sigma)$,$v = r+s$,有 $(\rho+\sigma)(v) = (\rho+\sigma)(r)+(\rho+\sigma)(s) = r+s = v$,则 $v \in \mathrm{Im}(\rho+\sigma)$,所以

$$\mathrm{Im}(\rho+\sigma) = \mathrm{Im}(\rho) \oplus \mathrm{Im}(\sigma).$$

定理 8.6.3　设 $\rho, \sigma \in L(V)$ 是射影,其中 V 是特征 $\neq 2$ 的域上的向量空间,那么 $\rho+\sigma$ 是一个射影 $\Leftrightarrow \rho \perp \sigma$,这样,$\rho+\sigma$ 是 $\mathrm{Im}(\rho) \oplus \mathrm{Im}(\sigma)$ 上沿 $\mathrm{Ker}(\rho) \cap \mathrm{Ker}(\sigma)$ 的射影.

下面我们考虑差 $\rho-\sigma$. $\rho-\sigma$ 是射影 $\Leftrightarrow \theta = \iota-(\rho-\sigma) = (\iota-\rho)+\sigma$ 是射影.

因此,我们也可以把之前的相关定理运用到这种情况中,得出

$$\rho-\sigma \text{ 是射影} \Leftrightarrow (\iota-\rho)\sigma = \sigma(\iota-\rho) = 0.$$

或等价地,$\rho\sigma = \sigma\rho = 0$,且 $\rho-\sigma = \iota-\theta$ 是 $\mathrm{Ker}(\theta)$ 上沿 $\mathrm{Im}(\theta)$ 的射影. 因为 $\theta = (\iota-\rho)+\sigma$,所以

$$\mathrm{Im}(\theta) = \mathrm{Im}(\iota-\rho) \oplus \mathrm{Im}(\sigma) = \mathrm{Ker}(\rho) \oplus \mathrm{Im}(\sigma);$$

$$\mathrm{Ker}(\theta) = \mathrm{Ker}(\iota-\rho) \cap \mathrm{Ker}(\sigma) = \mathrm{Im}(\rho) \cap \mathrm{Ker}(\sigma).$$

我们就证明了下面的

定理 8.6.4　设 $\rho, \sigma \in L(V)$ 是射影,其中 V 是特征 $\neq 2$ 的域上的向量空间,那么 $\rho-\sigma$ 是射影 $\Leftrightarrow \rho\sigma = \sigma\rho = \sigma$. 这里 $\rho-\sigma$ 是 $\mathrm{Im}(\rho) \cap \mathrm{Ker}(\sigma)$ 上沿 $\mathrm{Ker}(\rho) \oplus \mathrm{Im}(\sigma)$ 的射影.

定理 8.6.5　设 $\rho, \sigma \in L(V)$ 是射影. 如果 $\rho\sigma = \sigma\rho$,那么 $\rho\sigma$ 是一个射影,$\rho\sigma$ 是 $\mathrm{Im}(\rho) \cap \mathrm{Ker}(\sigma)$ 上沿 $\mathrm{Ker}(\rho)+\mathrm{Ker}(\sigma)$ 的射影.

证明　若 $\rho\sigma = \sigma\rho$,那么 $(\rho\sigma)^2 = \rho\sigma\rho\sigma = \rho^2\sigma^2 = \rho\sigma$,所以 $\rho\sigma$ 是一个射影. 若 $v = \rho\sigma(v)$,那么 $\rho(v) = \rho(\rho\sigma(v)) = \rho\sigma(v) = v$,所以 $v \in \mathrm{Im}(\rho)$. 同理,$v \in \mathrm{Im}(\sigma)$,所以

$$\mathrm{Im}(\rho\sigma) \subset \mathrm{Im}(\rho) \cap \mathrm{Im}(\sigma).$$

反包含是显然的,所以 $\mathrm{Im}(\rho\sigma) = \mathrm{Im}(\rho) \cap \mathrm{Im}(\sigma)$.

其次 ,注意到若 $v \in \mathrm{Ker}(\rho\sigma)$,那么 $\rho\sigma(v) = 0$,所以 $\sigma(v) \in \mathrm{Ker}(\rho)$. 因此 $v = \sigma(v)+(\iota-\sigma(v)) \in \mathrm{Ker}(\rho)+\mathrm{Ker}(\sigma)$. 而且,若 $v = r+s \in \mathrm{Ker}(\rho)+\mathrm{Ker}(\sigma)$,那么

$$\rho\sigma(v) = \rho\sigma(r+s) = \rho\sigma(r)+\rho\sigma(s) = 0+0 = 0.$$

所以 $v \in \mathrm{Ker}(\rho\sigma)$. 因此, $\mathrm{Ker}(\rho\sigma) = \mathrm{Ker}(\rho)+\mathrm{Ker}(\sigma)$.

如果 $\rho = \sigma$,那么 $\mathrm{Ker}(\rho) = \mathrm{Ker}(\sigma)$,所以上面的和不必是一直和.

8.6.3　射影的可对角化性

若 ρ 是射影,那么 $\rho \perp (\iota-\rho)$,$\rho+(\iota-\rho) = \iota$. 可把该式推广到两个以上的射影中.

定义 8.6.3　若 $\rho_1, \cdots, \rho_k$ 是射影，且

1）对于 $i \neq j, \rho_i \perp \rho_j$；　2）$\rho_1 + \cdots + \rho_k = \iota$，

那么我们把 2）中的和称为**单位分解**.

下面这个定理说明 V 的直和分解与单位分解之间的一一对应关系.

定理 8.6.6　1）若 $\rho_1 + \cdots + \rho_k = \iota$ 是单位分解，那么 $V = \mathrm{Im}(\rho_1) \oplus \cdots \oplus \mathrm{Im}(\rho_k)$.

2）反之，若 $V = S_1 \oplus \cdots \oplus S_k$，并且 ρ_i 是 S_i 上沿 $S_1 \oplus \cdots \oplus \hat{S}_i \oplus \cdots \oplus S_k$ 的射影，这里^指的是其对应项从直和中去掉，那么 $\rho_1 + \cdots + \rho_k = \iota$ 是单位分解.

证明　1）设 $\rho_1 + \cdots + \rho_k = \iota$ 是单位分解，对 $\forall v \in V$，$v = \iota v = \rho_1(v) + \cdots + \rho_k(v)$. 这说明 $V = \mathrm{Im}(\rho_1) + \cdots + \mathrm{Im}(\rho_k)$. 因为对于任意 i，射影 ρ_i 是正射影，所以对于 $\forall j \neq i$，我们有 $\mathrm{Im}(\rho_j) \subset \mathrm{Ker}(\rho_i)$. 因此 $\sum_{j \neq i} \mathrm{Im}(\rho_j) \subset \mathrm{Ker}(\rho_i)$，即 $\mathrm{Im}(\rho_i) \cap \sum_{j \neq i} \mathrm{Im}(\rho_j) = \{0\}$，所以 $V = \mathrm{Im}(\rho_1) \oplus \cdots \oplus \mathrm{Im}(\rho_k)$.

2）对于 $i \neq j, \mathrm{Im}(\rho_i) = S_i \subset \mathrm{Ker}(\rho_j)$，所以 $\rho_i \perp \rho_j$. 同样，

$$\iota v = s_1 + \cdots + s_k = \rho_1(v) + \cdots + \rho_k(v) = \sum_u \rho_u(v),$$

所以 $\iota = \rho_1 + \cdots + \rho_k$ 是一个单位分解.

现在我们考虑线性组合 $\tau = \lambda_1 \rho_1 + \cdots + \lambda_k \rho_k$，其中 $\rho_1 + \cdots + \rho_k = \iota$ 是一个单位分解. 因为任意向量 $v \in V$ 都有形式 $v = s_1 + \cdots + s_k$，其中 $s_i \in \mathrm{Im}(\rho_i)$，所以

$$\tau(v) = \sum_i \lambda_i \rho_i(v) = \sum_i \lambda_i s_i.$$

因此，τ 的作用可以用特别简单的形式表示出来：

$$v = s_1 + \cdots + s_k \Rightarrow \tau(v) = \lambda_1 s_1 + \cdots + \lambda_k s_k.$$

定理 8.6.7　线性算子 $\tau \in L(V)$ 是可对角化算子 $\Leftrightarrow$ 对于任意单位分解 $\rho_1 + \cdots + \rho_k = \iota$，$\tau$ 有形式 $\tau = \lambda_1 \rho_1 + \cdots + \lambda_k \rho_k$. 而且，若 τ 有上面形式，其中 λ_i 互异，ρ_i 非零，那么 λ_i 是 τ 的本征值，$\mathrm{Im}(\rho_i)$ 是 τ 的属于本征值 λ_i 的本征空间.

证明　设 τ 可对角化，$\lambda_1, \cdots, \lambda_k$ 是 τ 的互异本征值，那么 $V = \delta_{\lambda_1} \oplus \cdots \oplus \delta_{\lambda_k}$.

因为 δ_{λ_i} 上沿 $\delta_{\lambda_1} \oplus \cdots \oplus \hat{\delta}_{\lambda_i} \oplus \cdots \oplus \delta_{\lambda_k}$ 的射影 ρ_i 构成一个单位分解，而且若 $v \in \delta_{\lambda_i}$，则 $\tau(v) = \lambda_i v = (\lambda_1 \rho_1 + \cdots + \lambda_k \rho_k)(v)$，从而得出 $\tau = \lambda_1 \rho_1 + \cdots + \lambda_k \rho_k$.

反之，设 $\tau = \lambda_1 \rho_1 + \cdots + \lambda_k \rho_k$，利用定理 8.6.3，可以假定 λ_i 是互异的. 同样，$V = \mathrm{Im}(\rho_1) \oplus \cdots \oplus \mathrm{Im}(\rho_k)$. 由 $v = s_1 + \cdots + s_k \Rightarrow \tau(v) = \lambda_1 s_1 + \cdots + \lambda_k s_k$.

对于 $\forall s_i \in \mathrm{Im}(\rho_i)$，都有 $\tau(s_i) = \lambda_i s_i$，这说明 λ_i 是 τ 的本征值，且 $\mathrm{Im}(\rho_i) \subset \delta_{\lambda_i}$. 为了说明反包含，设 $v = s_1 + \cdots + s_k \in \delta_{\lambda_i}$，其中 $s_i \in \mathrm{Im}(\rho_i)$，那么

$$\lambda_i(s_1 + \cdots + s_k) = \lambda_i v = \tau(v) = \lambda_1 s_1 + \cdots + \lambda_k s_k,$$

则对于 $\forall j$，$(\lambda_i-\lambda_j)s_j=0$，因为 λ_j 互异，所以对于 $j\neq i$，$s_j=0$. 因此，$v=s_j\in \mathrm{Im}(\rho_i)$，从而 $\mathrm{Im}(\rho_i)=\delta_{\lambda_i}$，这说明 $V=\delta_{\lambda_1}\oplus\cdots\oplus\delta_{\lambda_k}$. 所以 τ 是可对角化的.

最后，注意到若 λ 是 τ 的本征值，且有本征向量 $v=s_1+\cdots+s_k$，则

$$\lambda(s_1+\cdots+s_k)=\lambda v=\tau(v)=\lambda_1 s_1+\cdots+\lambda_k s_k.$$

所以对于 $\forall i$，$(\lambda-\lambda_i)s_i=0$. 因此，若对于 $\forall i$，$\lambda\neq\lambda_i$，那么对于 $\forall i$，$s_i=0$，即 $v=0$，而这是不成立的. 从而，任意本征值都为系数 λ_i.

定义 8.6.4　有限维向量空间 V 上线性算子的本征值所成的集合称为 τ 的**谱**. 若 τ 是可对角化的，且 $\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k$，其中 $\rho_1+\cdots+\rho_k=\iota$ 是一个单位分解，λ_i 互异，ρ_i 非零，那么称为算子 τ 的**谱分解**.

8.6.4　射影与不变性

射影与线性算子 τ 的不变子空间及约化与这两个概念有关.

定理 8.6.8　设 $\tau\in L(V)$.

1）如果 S 是 τ 之下的一个不变子空间，那么对 S 上任意射影 ρ，$\rho\tau\rho=\tau\rho$.

2）如果 S 是 V 的一个子空间，且对 S 上任意射影 ρ，$\rho\tau\rho=\tau\rho$，那么 S 是 τ 下的一个不变子空间.

证明　1）设 S 是 τ 下的不变子空间，且设 ρ 是 S 上沿 T 的射影，从而 $V=S\oplus T$. 现在设 $v=s+t\in V$，其中 $s\in S$，$t\in T$，因为 ρ 固定 S，$\rho\tau\rho(v)=\rho\tau(s)=\tau(s)=\tau\rho(v)$，所以 $\rho\tau\rho=\tau\rho$.

2）设 $\rho\tau\rho=\tau\rho$，其中 ρ 是 S 上沿 T 的射影，设 $s\in S$，因为 ρ 固定 S，$\rho\tau(s)=\rho\tau\rho(s)=\tau\rho(s)=\tau(s)$，所以 ρ 固定 $\tau(s)$，则 $\tau(s)\in S$，S 在 τ 下是不变子空间.

定理 8.6.9　设 $V=S\oplus T$，那么线性算子 $\tau\in(V)$ 由 (S,T) 约化，当且仅当 $\tau\rho=\rho\tau$，其中 ρ 是 S 上沿 T 的射影.

第9章 赋范线性空间

本章在线性变换、欧氏空间和二次型的基础上，继续讨论线性泛函、实(复)内积空间、赋范空间和度量空间等内容，并给出傅立叶展开和施密特正交化方法.

9.1 线性泛函

9.1.1 线性泛函与对偶空间.

从 V 到基域 F 上的线性变换非常重要.

定义 9.1.1 令 V 是 F 上一向量空间. 值在基域 F 上的线性变换 $f \in L(V, F)$ 叫做 V 上的一**线性泛函**(简称为**泛函**). V 上全体线性泛函所成的集合用 V^* 表示，叫做 V 的**代数对偶空间**.

因为存在另一种定义在赋范向量空间上的对偶空间，其线性变换的连续性是有意义的. 我们将在后面讨论连续对偶空间.

为了便于区分线性泛函和其他类型的线性变换，我们通常用小写罗马字母，如 f, g, h 来表示线性泛函.

例 9.1.1 由 $f(p(x)) = p(0)$ 定义的映射 f: $F[x] \to F$ 是一线性泛函，且在 0 求值.

例 9.1.2 令 $C[a, b]$ 表示 $[a, b]$ 上全体连续实函数所成的向量空间，且令 f: $C[a, b] \to \mathbf{R}$ 由 $f(\alpha(x)) = \int_a^b \alpha(x)\mathrm{d}(x)$ 定义，那么 $f \in C[a, b]^*$. 对于 $\forall f \in V^*$, $\dim(\mathrm{Ker}(f)) + \dim(\mathrm{Im}(f)) = \dim(V)$.

但是，因为 $\mathrm{Im}(f) \subset F$，所以 $\mathrm{Im}(f) = \{0\}$，在这一情形下，f 为零线性泛函；或 $\mathrm{Im}(f) = F$，在这一情形下，f 为满射的. 也就是说，非零线性泛函是满射的，而且，如果 $\dim(V) < \infty$，那么

$$\dim(\mathrm{Ker}(f)) = \dim(V) - 1.$$

因此，线性泛函的核是域 V 的一个相对"大的"的子空间. 即使 V 是无限维的，$\mathrm{Ker}(f)$也有余维数 0 或 1.

定理 9.1.1 1) 对于任意非零向量 $v \in V$，存在线性泛函 $f \in V^*$，使得 $f(v) \neq 0$.

2）向量 $v \in V = 0$，当且仅当对于任意 $f \in V^*$，$f(v) = 0$.

例 9.1.3 对于任意非零向量 $v \in V$，存在一线性泛函 $f \in V^*$，使得 $f(v) \neq 0$.

设 V 是有限维的，令 $B = \{v_1, \cdots, v_n\}$ 是 V 的一个基. 对于每一 $1 \leqslant i \leqslant n$，由正交性条件有 $f_i(v_j) = \delta_{i,j}$，$j = 1, \cdots, n$.

可以定义一个线性泛函 $f_i \in V^*$，其中 $\delta_{i,j}$ 被认为是克罗内克 δ 函数，定义为

$$\delta_{i,j} = \begin{cases} 1 & 若\ i = j; \\ 0 & 若\ i \neq j. \end{cases}$$

定理 9.1.2 令 $B = \{v_1, \cdots, v_n\}$ 是 V 的一个基，那么由 $f_i(v_j) = \delta_{i,j}(j = 1, \cdots, n)$ 定义的线性泛函 $f_1, \cdots, f_n$ 作成对偶空间 V^* 的一个基，基 $B^* = \{f_1, \cdots, f_n\}$ 叫做 B 的**对偶基**.

证明 如果 $0 = r_1 f_1 + \cdots + r_n f_n$，其中 0 代表零线性泛函，那么我们可以把这一等式的两边都应用到基向量 v_i 上，得到

$$0 = 0(v_i) = \sum_j r_j f_j(v_i) = \sum_j r_j \delta_{i,j} = r_i.$$

所以对于 $\forall i, r_i = 0$. 因此，B^* 是线性无关的. $\forall \xi \in V^*$ 都由它在基向量 v_i 上的值唯一确定，即 $\xi(v_i) = a_i$. 令 η 为线性泛函 $\eta = a_1 f_1 + \cdots + a_n f_n$，那么 $\eta(v_j) = a_j = \xi(v_j)$. 所以 $\xi = \eta \in \mathrm{span}\{f_1, \cdots, f_n\}$，从而 B^* 生成 V^*. 因此，B^* 是 V^* 的一个基.

推论 9.1.3 若 $\dim(V) < \infty$，则 $\dim(V^*) = \dim(V)$.

下面的例子说明没有有限性条件，上述推论不成立.

例 9.1.4 令 V 是域 $F = \mathbf{Z}_2 = \{0,1\}$ 上一无限维向量空间，且有基 B. 由于 F 中仅有系数 0 和 1，所以 F 上一有限线性组合正好是一个有限和. 因此，V 是 B 中向量的全体有限和所成的集合，从而 $|V| = |P_0(B)| = |B|$.

另一方面，每一线性泛函 $\xi \in V^*$ 是通过取定 ξ 在基 B 上的值来唯一确定的. 由于这些值一定为 0 或 1，因此取定一个线性泛函相当于取定 B 的子集，其中 ξ 的取值为 1. 换句话说，V 上的线性泛函和 B 的所有子集之间一一对应. 因此，

$$|V^*| = |P(B)| > |B| = |V|.$$

这说明 V^* 既不能同构于 V，也不能同构于 V 的任意一个真子集. 因此，

$$\dim(V^*) > \dim(V).$$

例 9.1.5 令 S 是有限维向量空间 V 的真子空间，设 $v \in V - S$，则存在一线性泛函 $f \in V^*$，使得对于 $\forall s \in S, f(v) = 1, f(s) = 0$.

9.1.2 自反性

如果 V 是一向量空间，那么对偶空间 V^* 也是一个向量空间. 因此，我们可以作成

双对偶空间 V^{**}，它由全体线性泛函Σ：$V^* \to F$组成. 换句话说，V^{**}的元素Σ是一线性映射，这一线性映射把标量分配到V的每一线性泛函上.

这样，我们就有一种相当简便的方法来得到V^{**}中的元素，即，如果$v \in V$，考虑由$\overline{v}(f) = f(v)$定义的映射$\overline{v}$：$V^* \to F$，$\overline{v}$把线性泛函f映到标量$f(v)$上. 这一映射叫做在v求值. 为了说明$\overline{v}$在V^{**}中，我们必须证明它是线性的. 如果$f, g \in V^*$，那么

$$\overline{v}(rf+sg) = (rf+sg)(v) = rf(v)+sg(v) = r\overline{v}(f)+s\overline{v}(g),$$

所以$\overline{v}$确实是线性的.

定义 9.1.2　$\forall v \in V$，对v求值都在V^{**}中，所以由$\tau(v) = \overline{v}$定义映射τ：$V \to V^{**}$，τ叫做从V到V^{**}的**自然映射**.

τ是单射，而且在有限维的情形下，它又是满射的.

定理 9.1.4　通过$\tau(v)$在v求值定义的典范映射τ：$V \to V^{**}$是一个单一同态，而且，如果V是有限维的，那么τ是一同构.

证明　为了说明τ是线性的，我们注意到$\tau(ru+sv) = \overline{ru+sv}$是在$ru+sv$求值. 但是对于$\forall f \in V^*$，$\overline{ru+sv}(f) = f(ru+sv) = rf(u)+sf(v) = (r\overline{u}+s\overline{v})(f)$. 因此$\tau(ru+sv) = \overline{ru+sv} = r\overline{u}+s\overline{v} = r\tau(u)+s\tau(v)$，这说明$\tau$是线性的.

为了确定τ的核，我们有$\tau(v) = 0 \Rightarrow \overline{v} = 0 \Rightarrow \overline{v}(f) = 0$，对于$\forall f \in V^* \Rightarrow f(v) = 0$，对于$\forall f \in V^* \Rightarrow v = 0$，所以$\mathrm{Ker}(\tau) = \{0\}$，即$\tau$是单射的.

在有限维的情形下，我们有$\dim(V^{**}) = \dim(V^*) = \dim(V)$，所以$\tau$也是满射的，从而$\tau$是一个同构.

若$\dim(V) < \infty$，由于V和V^{**}的维数相同，所以$V \cong V^{**}$. 自然映射$v \to \overline{v}$是一同构，所以V称为是**代数自反**的. 因此，一切有限维向量空间都是代数自反的.

如果V是有限维的，我们把双对偶空间V^{**}和V看成是等价的，并把V^{**}的元素看成是V的向量.

以下例子表明，向量空间并不是代数自反的.

例 9.1.6　令V是$F = \{0,1\}$上一向量空间，且有可数个无限有序基$B = (b_1, b_2, \cdots)$，那么任意向量$v \in V$都与它的坐标序列$v = (a_1, \cdots, a_n)$等价. 这里$a_i \in \{0,1\}$，且仅含有有限个a_i等于1.

另一方面，任意$f \in V^*$由它在B中的向量的值唯一确定，而且，由于这些值可以任意选取，所以可以认为f与二元序列$f = (\alpha_1, \alpha_2, \cdots)$等价，且序列中1的数量没有限制.

现在，我们定义二元序列$x = (x_1, x_2, \cdots)$的**支集**，记作$\mathrm{supp}(x)$，为坐标位置所成的集合，其中$x_i = 1$，因此V中的向量是一个含有有限支集的二元序列，而V上的线性泛函是任意二元序列.

由 f 的表示可以得出 $\overline{v}(f) = f(v) = |\operatorname{supp}(v) \cap \operatorname{supp}(f)|$. 对于 $\forall v \in V$,通过找出线性泛函 $\varphi \in V^{**}$,这里 φ 没有形式 $\overline{v}$,可以证明典范映射 $\tau: v \to \overline{v}$ 不是满射的.

最后,由 $e_k = (0,\cdots,0,1,0,\cdots)$ 定义线性泛函 $e_k \in V^*$,这里 1 在第 k 个位置,那么 $\operatorname{supp}(e_k) = \{k\}$,所以 $\overline{v}(e_k) = |\operatorname{supp}(v) \cap \operatorname{supp}(e_k)| = |\operatorname{supp}(v) \cap \{k\}|$. 因此,对于 $\forall v \in V$,映射 $\overline{v}$ 有性质 $k > \max\{\operatorname{supp}(\overline{v})\} \Rightarrow \overline{v}(e_k) = 0$.

现在,由于线性泛函 e_k 线性无关,因此我们可以把集合 $\{e_k\}$ 扩充到 V^* 的基 B 上. 对于 $\forall k$,通过规定 $\varphi(e_k) = 1$,再把 φ 任意地扩充到基 B 的所有向量上,我们可以定义映射 $\varphi \in V^{**}$(即,对于 φ 如何定义在 B 的其他元素上,我们忽略不计),这样,φ 在 V^{**} 上定义了一个线性泛函,对于 $\forall k$,有性质

$$\varphi(e_k) = 1.$$

但上式与 $k > \max\{\operatorname{supp}(\overline{v})\} \Rightarrow \overline{v}(e_k) = 0$ 一起说明,对于 $\forall v \in V$,φ 不存在形式 $\overline{v}$. 因此,典范映射不是满射的,而且 V 不是代数自反的.

9.1.3 零化子

如果 $f \in V^*$,那么 f 定义在 V 中的向量上. 通过令 $f(M) = \{f(v) \mid v \in M\}$ 我们也可以在 V 的子集 M 上定义 f.

定义 9.1.3 令 M 是向量空间 V 的一非空子集. M 的零化子 M^0 为

$$M^0 = \{ f \in V^* \mid f(M) = 0\}.$$

"零化子"具有描述性,因为 M^0 由 M 中把每一向量映成 0 的所有线性泛函组成.

即使 M 不是 V^* 的子空间,M^0 仍是 V^* 的一子空间,这样我们证明了以下

定理 9.1.5 W 是有限维向量空间 V 的子空间,那么 $\dim(W^0) = \dim(V) - \dim(W)$.

证明 令 $\{u_1, \cdots, u_k\}$ 是 W 的一个基,并把它扩充到 V 的基

$$B = \{u_1, \cdots, u_k, v_1, \cdots, v_{n-k}\}$$

上. 令 $B^* = \{\mu_1, \cdots, \mu_k, v_1, \cdots, v_{n-k}\}$ 是 B 的对偶基.

我们要说明 $\{v_1, \cdots, v_{n-k}\}$ 是 W^0 的一个基. 由于这一集合线性无关,因此我们只要说明它生成 W^0. 但如果 $f \in W^0$,那么由于 $f \in V^*$,所以

$$f = r_1\mu_1 + \cdots + r_k\mu_k + s_1 v_1 + \cdots + s_{n-k} v_{n-k}.$$

由于对于 $\forall v \in W$, $f(v) = 0$,所以对于 $i = 1, \cdots, k$, $0 = f(u_i) = r_i$. 因此,$f = s_1 v_1 + \cdots + s_{n-k} v_{n-k}$,这说明 $\{v_1, \cdots, v_{n-k}\}$ 生成 S^0.

例 9.1.7 在无限维的情形下,V 是 $F = \{0,1\}$ 上的向量空间,这一向量空间有可数个无限有序基 $B = (b_1, b_2, \cdots)$. 令 S 是 V 的含有有序基 $X = (b_1)$ 的子空间,T 是 V 的含有有序基 $\Delta = (b_2, b_3, \cdots)$ 的子空间. 因为 $|\Delta| = |B|$,所以 $T \cong V$.

现在考虑零化子 W^0. 通过规定 $\hat{f}(b_1) = 0$,任意线性的 $f \in T^*$ 都可以扩充到线性泛函 $\hat{f} \in V^*$ 上. 而且,对于 $\forall f \in T^*$, W^0 中任意线性泛函都有形式 $\hat{f}$,因为 $f \in W^0$, $f(b_1) = 0$,所以 W^0 和 T^* 一一对应,从而 $|W^0| = |T^*|$. 又由 $T \cong V$ 推出 $T^* \cong V^*$,所以 $|T^*| = |V^*|$,有 $|W^0| = |V^*|$. 又因为 $|V^*| > |V|$,所以 $|W^0| > |V|$,则 W^0 不能同构于 V 或 V 的任意一个子空间. 因此,$\dim(W^0) > \dim(V)$.

零化子的基本性质如下:

定理 9.1.6　1) 对于 V 的任意子集 M 和 N, $M \subset N \Rightarrow N^0 \subset M^0$.

2) 如果 $\dim(V) < \infty$,那么在自然映射之下 V^{**} 与 V 是等价的,即有 $M^{00} = \mathrm{span}(M)$. 特别地,若 W 是 V 的一子空间,那么 $W^{00} = W$.

3) 如果 $\dim(V) < \infty$, W 和 T 是 V 的子空间,那么

$$(W \cap T)^0 = W^0 + T^0,\ (W + T)^0 = W^0 \cap T^0.$$

定理 9.1.7　令 $V = W \oplus T$,那么

1) $W^* \cong T^0$, $T^* \cong W^0$.

2) $(W \oplus T)^* = W^0 \oplus T^0$.

证明　首先,我们证 $W^* \cong T^0$. 这表明 V 上零化直和项 T 的任意一个线性泛函只是其补子空间 W 上的一个线性泛函,而这显然是合理的.

令 $f \in T^0 \subset V^*$,因此 $f(T) = 0$. 映射 τ: $f \to f|_S$ 是线性的,这里 τ 从泛函 $f \in V^*$ 映到限制 $f|_W$ 上,且限制 $f|_W$ 在 W^* 中. 而且,如果 $f|_W = 0$,那么 $f(W) = 0$,由于 $f(T) = 0$,所以 $f = 0$. 因此 τ 是单射的.

我们要说明 τ 是满射的,即,对于 $g \in W^*$,我们必须找出一个 $f \in T^0$,使得对于 $\forall s \in W$, $f|_W(s) = g(s)$. 换句话说,我们想要把 g"扩充"到 V 上,这样就扩充在 T^0 中了. 所以,我们只要在 T 上定义扩充为 0.

特别地,令 $f \in V^*$ 由 $f(s+t) = g(s)$ 定义,那么 f 是唯一确定和线性的,而且 $f \in T^0$. 因为对于 $\forall t \in T$, $f(t) = f(0+t) = g(0) = 0$,所以 $f|_W$ 的确是 g,所以 τ 是一同构, $T^0 \cong W^*$. 根据对称性,也有 $W^0 \cong T^*$.

2) 令 $f \in W^0 \cap T^0$,那么 $f(W) = 0 = f(T)$,则 $f = 0$. 因此, $W^0 \cap T^0 = \{0\}$. 由于 W^0 和 T^0 是 V^* 的子空间,所以我们有 $(W \oplus T)^* \supset W^0 \oplus T^0$.

另一方面,如果 $f \in (W \oplus T)^*$,那么由 $g(s+t) = f(t)$ 和 $h(s+t) = f(s)$,定义 $g, h \in (W \oplus T)^*$. 易知这些映射是唯一确定和线性的,而且 $g(W) = 0$, $h(T) = 0$,所以 $g \in W^0$, $h \in T^0$.

最后, $f(s+t) = f(t) + f(s) = g(s+t) + h(s+t) = (g+h)(s+t)$,从而 $f = g + h \in W^0 \oplus T^0$. 因此, $(W \oplus T)^* \subset W^0 \oplus T^0$,获证.

9.1.4 伴随算子

定义 9.1.4 如果 $\tau \in L(V, W)$，那么对于 $f \in W^*$，由 $\tau^*(f) = f \circ \tau = f\tau$，我们可以定义映射 $\tau^*: W^* \to V^*$. 这是成立的，因为 $\tau: V \to W$ 和 $f: W \to F$，所以复合 $f\tau: V \to F$ 在 V^* 中. 因此，$\forall v \in V, \tau^*(f)(v) = f(\tau(v))$. 映射 τ^* 叫做 τ 的**伴随算子**.

以下是伴随算子的基本性质.

定理 9.1.8 1) $(\tau + \sigma)^* = \tau^* + \sigma^*$，对于 $\tau, \sigma \in L(V, W)$.

2) $(r\tau)^* = r\tau^*$，对于任意 $r \in F$ 和 $\tau \in L(V, W)$.

3) $(\tau\sigma)^* = \sigma^*\tau^*$，对于 $\tau \in L(V, W)$ 和 $\sigma \in L(W, U)$.

4) $(\tau^{-1})^* = (\tau^*)^{-1}$，对于任意可逆 $\tau \in L(V)$.

证明 我们只证明论断 3) 和 4). 对于 $\forall f \in U^*$，论断 3) 从以下事实中得到

$$(\tau\sigma)^*(f) = f\tau\sigma = \sigma^*(f\tau) = \tau^*(\sigma^*(f)) = (\tau^*\sigma^*)(f).$$

论断 4) 从下面的事实中得到 $\tau^*(\tau^{-1})^* = (\tau^{-1}\tau)^* = \iota^* = \iota$. 同理，$(\tau^{-1})^*\tau^* = \iota$.

若 $\tau \in L(V, W)$，那么 $\tau^* \in L(W^*, V^*)$，我们可以构造 $\tau^{**} \in L(V^{**}, W^{**})$. 当然，$\tau^{**}$ 不等于 τ. 然而，在有限维的情形下，如果我们用自然映射分别把 V^{**} 和 V，W 和 W^{**} 看成等价，那么 τ^{**} 就在 $L(V, W)$ 中，从而 τ^{**} 等于 τ.

定理 9.1.9 令 V 为有限维的，设 $\tau \in L(V, W)$. 如果我们用自然映射把 V^{**} 和 V，W 和 W^{**} 看成等价，那么 $\tau^{**} = \tau$.

证明 我们有 $\tau^{**}: V^{**} \to W^{**}$，且对于 $\forall f \in W^*$，

$$\tau^{**}(\overline{v})(f) = \overline{v}\tau^*(f) = \overline{v}(f\tau) = f\tau(v) = \overline{\tau(v)}(f).$$

所以 $\tau^{**}(\overline{v}) = \overline{\tau(v)}$. 因此，对于 $\forall v \in V$，有，$\tau^{**}(v) = \tau(v)$，所以 $\tau^{**} = \tau$.

定理 9.1.10 令 $\tau \in L(V, W)$，那么

1) $\mathrm{Ker}(\tau^*) = \mathrm{Im}(\tau)^0$.

2) $\mathrm{Im}(\tau^*)^0 = \mathrm{Ker}(\tau)$.

3) $\mathrm{Im}(\tau^*) \subset \mathrm{Ker}(\tau)^0$.

4) 若 $\dim(V) < \infty, \dim(W) < \infty$，那么 $\mathrm{Im}(\tau^*) = \mathrm{Ker}(\tau)^0$.

证明 $\tau: V \to W, \tau^*: W^* \to V^*$.

1) 我们有

$$f \in \mathrm{Ker}(\tau^*) \Leftrightarrow \tau^*(f) = 0 \Leftrightarrow f\tau = 0 \Leftrightarrow \text{对于 } \forall v \in V,$$

$$f(\tau(v)) = 0 \Leftrightarrow f(\mathrm{Im}(\tau)) = 0 \Leftrightarrow f \in \mathrm{Im}(\tau)^0.$$

2) 我们有

$$v \in \mathrm{Ker}(\tau) \Leftrightarrow \tau(v) = 0 \Leftrightarrow \text{对于 } \forall f \in W^*, f(\tau(v)) = 0$$

$$\Leftrightarrow \text{对于 } \forall f \in W^*, \tau^*(f)(v) = 0$$

$$\Leftrightarrow \text{对于 } \forall f \in W^*, \bar{v}(\tau^*(f)) = 0$$

$$\Leftrightarrow \bar{v} \in \mathrm{Im}(\tau^*)^0.$$

3) 对 $\forall v \in \mathrm{Ker}(\tau)$ 和 $f \in W^*$，$\tau^*(f)(v) = f(\tau(v)) = 0$，则 $\tau^*(f)(\mathrm{Ker}(\tau)) = 0$，即 $\tau^*(f) \in \mathrm{Ker}(\tau)^0$，由于对于 $\forall f \in W^*$，以上也成立，因此 $\mathrm{Im}(\tau^*) \subset \mathrm{Ker}(\tau)^0$.

4) 当向量空间为有限维时，由论断2)可得 $\mathrm{Im}(\tau^*) \subseteqq \mathrm{Im}(\tau^*)^{00} \subseteqq \mathrm{Ker}(\tau)^0$. 但是，根据论断3)，$\mathrm{Im}(\tau^*) \subset \mathrm{Ker}(\tau)^0$，所以这些空间必相等.

令 $\tau \in L(V, W)$，V 和 W 有限维，那么 $rk(\tau) = rk(\tau^*)$.

在有限维的情形下，$\tau \in L(V, W)$ 和 $\tau^* \in L(W^*, V^*)$ 都可以用矩阵表示. 为了找出这两个矩阵之间的关系，设 $B = (b_1, \cdots, b_n)$，$C = (c_1, \cdots, c_m)$ 分别为 V 和 W 的有序基，且 $B^* = (b_1^*, \cdots, b'_n{}^*)$，$C^* = (c_1^*, \cdots, c_m^*)$ 为相应的对偶基. 如果令 $(\tau)_{B,C} = (\alpha_{i,j})$，那么在 $\tau(b_j)$ 中，$\alpha_{i,j}$ 是 c_i 的坐标.

另一方面，如果 $(\tau^*)_{C^*,B^*} = (\alpha_{i,j})$，那么在 $\tau^*(c_j^*)$ 中，$\alpha_{i,j}$ 是 b_i^* 的坐标. 但这一坐标是 $\tau^*(c_j^*)(b_i) = c_j^*(\tau(b_i))$. 在 $\tau(b_i)$ 中，$\alpha_{i,j}$ 是 c_j 的坐标，而这个坐标又是 $\alpha_{j,i}$，即 $\alpha_{i,j} = \alpha_{j,i}$，所以

$$(\tau^*)_{C^*,B^*} = ((\tau)_{B,C})^{\mathrm{T}}.$$

定理 9.1.11 令 $\tau \in L(V, W)$，其中 V 和 W 是有限维的. 如果 B 是 V 的有序基，C 是 W 的有序基，C^* 和 X^* 为相应的对偶基，那么

$$(\tau^*)_{C^*,B^*} = ((\tau)_{B,C})^{\mathrm{T}}.$$

9.2 内积空间

9.2.1 实(复)内积空间的定义

现在讨论实向量空间和复向量空间，在它们上面定义了一个叫做内积的附加函数. F 表示实数域或复数域. 若 α 是一个复数，那么 $\bar{\alpha}$ 就表示 α 的复共轭数.

定义 9.2.1 令 V 是 F 上的向量空间，这里 $F = \mathbf{R}$ 或 $F = \mathbf{C}$. V 上的内积是函数 $<,>: V \times V \to F$，并且有以下性质：

1)(正定性) 对于 $\forall v \in V$，$< v, v > \geqslant 0$ 且 $< v, v > = 0 \Leftrightarrow v = 0$.

2)(线性性) 对于 $\forall u, v \in V, r, s \in F$，$< ru + sv, w > = r < u, w > + s < v, w >$.

3) 对于 $F = \mathbf{C}$(共轭对称)：$< u, v > = \overline{< v, u >}$.

对于 $F = \mathbf{R}$(对称性)：$< u, v > = < v, u >$.

实(或复)向量空间 V,加之 V 上定义的内积,叫做**实(或复)内积空间**.

性质 1)说明数 $\langle v, v\rangle$ 总是实数,即使 V 是复向量空间. 把性质 2)和 3)结合起来,在复数域的情形下,我们得到

$$\begin{aligned}\langle w, ru+sv\rangle &= \overline{\langle ru+sv, w\rangle} = \bar{r}\,\overline{\langle u, w\rangle} + \bar{s}\,\overline{\langle v, w\rangle} \\ &= \bar{r}\langle w, u\rangle + \bar{s}\langle w, v\rangle.\end{aligned}$$

这称为**共轭线性**. 因此,复内积在其第一坐标中是线性的,在其第二坐标中是共轭线性的. 也说内积是**半双线性的**.

在实数域中,内积在这两个坐标中都是线性的——这个性质称为**双线性**.

例 9.2.1 1) 向量空间 $\mathbf{R}^n$ 在标准内积或点积下,由

$$\langle (r_1, \cdots, r_n), (s_1, \cdots, s_n)\rangle = r_1 s_1 + \cdots + r_n s_n$$

定义,是一个内积空间. 通常称内积空间 $\mathbf{R}^n$ 为 **n 维欧几里得空间**.

2) 向量空间 $\mathbf{C}^n$ 在标准内积下,由

$$\langle (r_1, \cdots, r_n), (s_1, \cdots, s_n)\rangle = r_1\bar{s}_1 + \cdots + r_n\bar{s}_n$$

定义,是一个内积空间. 通常称这一内积空间为 **n 维酉空间**.

3) 全体二元 n 元组所成的向量空间 $V(n, 2)$,在内积

$$\langle (r_1, \cdots, r_n), (s_1, \cdots, s_n)\rangle = (r_1 s_1 + \cdots + r_n s_n) \bmod 2$$

下,是一内积空间.

4) 闭区间 $[a, b]$ 上所有连续复值函数所成向量空间 $C[a, b]$,在内积

$$\langle f, g\rangle = \int_a^b f(x)\,\overline{g(x)}\,\mathrm{d}x$$

下,是一内积空间.

例 9.2.2 最重要的内积空间之一是在内积

$$\langle (s_n), (t_n)\rangle = \sum_{n=0}^{\infty} s_n \bar{t}_n$$

下,有性质 $\sum\limits_{n=0}^{\infty} |s_n|^2 < \infty$ 的所有复序列 (s_n) 所成的向量空间 L^2. 当然,为了使这一内积有意义,等式右边的和必须收敛.

为此,因为 $(s_n), (t_n) \in L^2$,

$$s = \sum_{n=0}^{\infty} |s_n|^2 < \infty, t = \sum_{n=0}^{\infty} |t_n|^2 < \infty;$$

又

$$0 \leqslant (|s_n| - |t_n|)^2 = |s_n|^2 - 2|s_n||t_n| + |t_n|^2,$$

所以 $2|s_n t_n| \leqslant |s_n|^2 + |t_n|^2$,从而

$$2\left|\sum_{n=0}^{\infty} s_n \bar{t}_n\right| \leqslant 2\sum_{n=0}^{\infty} |s_n t_n| \leqslant \sum_{n=0}^{\infty} |s_n|^2 + \sum_{n=0}^{\infty} |t_n|^2 = s + t < \infty.$$

易验证 L^2 是一个内积空间.

下面的结论虽然简单,但非常有用:

定理 9.2.1　如果 V 是一内积空间,且对于 $\forall x \in V, <u, x> = <v, x>$,则 $u = v$. 在 V 到 W 的内积的限制之下,内积空间 V 的向量子空间 W 也是一内积空间.

9.2.2　赋范线性空间

定义 9.2.2　如果 V 是一内积空间,那么由 $\| v \| = \sqrt{<v,v>}$,我们可以定义 $\forall v \in V$ 的**范数**,或**长度**. 向量 $v \in V$ 是一个**单位向量**,如果 $\| v \| = 1$.

下面是范数的基本性质:

定理 9.2.2　1) $\| v \| \geqslant 0$ 且 $\| v \| = 0 \Leftrightarrow v = 0$.

2) 对于 $\forall r \in F, v \in V, \| rv \| = | r | \| v \|$.

3)(柯西 — 施瓦兹不等式)　对于 $\forall u, v \in V, | <u, v> | \leqslant \| u \| \| v \|$ 等号成立,当且仅当对于 $\forall r \in F, u = rv$.

4)(三角不等式)　对于 $\forall u, v \in V, \| u + v \| \leqslant \| u \| + \| v \|$ 等号成立,当且仅当对于 $\forall r \in F, u = rv$.

5) 对于 $\forall u, v, x \in V, || u | - | v || \leqslant \| u - x \| + \| x - v \|$.

6) 对于 $\forall u, v \in V, || u | - | v || \leqslant \| u - v \|$.

7)(平行四边形法则)　对于 $\forall u, v \in V, \| u + v \|^2 + \| u - v \|^2 = 2\| u \|^2 + 2\| v \|^2$.

证明　我们只证论断 3)和 4).

3) 如果 u 或 v 为 0,获证. 设 $u, v \neq 0$,那么,对于 $\forall r \in F$,

$$\begin{aligned} 0 \leqslant \| u - rv \|^2 &= <u - rv, u - rv> \\ &= <u, u> - \bar{r} <u, v> - r[<v, u> - \bar{r} <v, v>]. \end{aligned}$$

(选取 $\bar{r} = <v,u>/<v,v>$,使方括号中的数为 0,从而得到)

$$\begin{aligned} &= <u, u> - \bar{r} <u, v> = <u, u> - \frac{<v,u><u,v>}{<v,v>} \\ &= \| u \|^2 - \frac{|<u,v>|^2}{\| u \|^2}. \end{aligned}$$

后面一个表达式的非负性等价于柯西—施瓦兹不等式. 而且,等式成立的充分必要

条件是 $\| u-rv \|^2=0$，即 $u-rv=0$，这说明 u 是 v 的数量倍数（如果 u 是 v 的数量倍数，那么很容易看出等式成立）.

4）我们利用柯西—施瓦兹不等式，得

$$\begin{aligned}\| u+v \|^2 &= \langle u+v, u+v\rangle = \langle u, u\rangle+\langle u, v\rangle+\langle v, u\rangle+\langle v, v\rangle \\ &\leqslant \| u \|^2+2\langle u, v\rangle+\| v \|^2 \\ &\leqslant \| u \|^2+2\| u \| \| v \|+\| v \|^2 \\ &= (\| u+v \|)^2.\end{aligned}$$

从中得出三角不等式 4).

定义 9.2.3　设 V 是一个向量空间，满足定理 9.2.2 中性质 1),2) 和 4) 的函数 $\| \ \|: V\to R$，称为一个**赋范线性空间**（函数 $\| \ \|$ 叫做**范数**）.

例 9.2.3　在 $\| v \| = \sqrt{\langle v,v\rangle}$ 给出的范数之下，任何内积空间都是赋范线性空间.

V 的内积可以由范数来表示.

定理 9.2.3　1）如果 V 是一个实内积空间，那么

$$\langle u, v\rangle = \frac{1}{4}(\| u+v \|^2-\| u-v \|^2).$$

2）如果 V 是一个复内积空间，那么

$$\langle u, v\rangle = \frac{1}{4}(\| u+v \|^2-\| u-v \|^2)+\frac{1}{4}\mathrm{i}(\| u+\mathrm{i}v \|^2-\| u-\mathrm{i}v \|^2).$$

上面的公式称为**极化恒等式**.

9.3　距离空间

9.3.1　距离空间

在内积空间中，范数可以用来定义任意两个向量之间的距离.

定义 9.3.1　令 V 是一个内积空间. 把 $d(u, v) = \| u-v \|$ 定义为 V 中任意两个向量 u 和 v 之间的距离 $d(u, v)$.

以下是距离的基本性质：

定理 9.3.1　1）$d(u, v)\geqslant 0$ 且 $d(u, v)=0$，当且仅当 $u=v$.

2）（对称性）　$d(u, v)=d(v, u)$.

3）（三角不等式）　$d(u, v)\leqslant d(u, w)+d(w, v)$.

定义 9.3.2　任何非空集 V，加之满足定理中性质的函数 $d: V\times V\to \mathbf{R}$，叫做距离

空间. 函数 d 叫做 V 的**距离**.

因此,在距离 $d(u, v) = \| u-v \|$ 之下,任何内积空间都是一个距离空间.

继续讨论之前,我们说明一下本章和下一章的主要内容——内积,然后是距离的出现产生了许多与收敛概念相关的拓扑问题.

定义 9.3.3　如果 $\lim\limits_{n\to\infty} d(v_n, v) = 0$,称内积空间中的向量序列 (v_n) **收敛**于 $v \in V$;或等价地,如果 $\lim\limits_{n\to\infty} \| v_n - v \| = 0$.

与收敛有关的一些比较重要的概念有封闭性和闭包、完全性,以及线性算子和线性泛函的连续性.

在有限维的情形下,情况就很简单——所有子空间都是闭的,所有内积空间都是完全的,所有线性算子和泛函都是连续的. 但是,在无限维的情形下,就没有这么简单了.

9.3.2　等距同构

我们要描述有限和无限维内积空间的一些基本性质,然后在有限维的情形下讨论几种特殊类型的算子(正规算子、酉算子和自伴算子).

向量空间的同构保持向量空间运算. 以下是内积空间的相应概念.

定义 9.3.4　令 V 和 W 是内积空间,且令 $\tau \in L(V, W)$,τ 是一个等距,如果它保持内积,即如果对于 $\forall u, v \in V$, $< \tau(u), \tau(v) > = < u, v >$.

双射等距叫做**等距同构**. 当 $\tau: V \to W$ 是一个同构映射时,我们就说 V 和 W 是等距同构的.

不难证明等距是单射,所以如果等距也是满射,那么它就是等距同构. 而且,如果 $\dim(V) = \dim(W) < \infty$,那么由单射性可得满射性,所以等距和等距同构的概念是等价的.

另一方面,下面的例子说明了在无限维的情形下,情况并非如此.

例 9.3.1　令 $\tau: L^\infty \to L^\infty$ 由 $\tau(x_1, x_2, x_3, \cdots) = (0, x_1, x_2, \cdots)$ 定义(这是右移算子),那么 τ 是一个等距,但显然 τ 不是满射.

定理 9.3.2　线性变换 $\tau \in L(V,W)$ 是一个等距,当且仅当 τ 保持范数,即,当且仅当对于 $\forall v \in V$, $\| \tau(v) \| = \| v \|$.

证明　显然,等距保持范数. 逆命题根据定理 9.2.3 得到. 在实内积空间中,如果 τ 保持了范数,那么

$$
\begin{aligned}
< \tau(u), \tau(v) > &= \frac{1}{4}(\| \tau(u) + \tau(v) \|^2 - \| \tau(u) - \tau(v) \|^2) \\
&= \frac{1}{4}(\| \tau(u+v) \|^2 - \| \tau(u-v) \|^2)
\end{aligned}
$$

$$= \frac{1}{4}(\| u+v \|^2 - \| u-v \|^2)$$
$$= < u, v >.$$

所以 τ 是一个等距. 复(实)内积空间中的证明类似.

下一个结论指出了实内积空间与复内积空间的主要不同点之一.

定理 9.3.3 令 V 是一内积空间,且令 $\tau \in L(V)$.

1) 如果对于 $\forall v, w \in V, < \tau(v), w > = 0$,那么 $\tau = 0$.

2) 如果 V 是复内积空间,并且对于 $\forall v \in V, < \tau(v), v > = 0$,那么 $\tau = 0$.

3) 对于实内积空间,2) 一般不成立.

证明 1) 可以直接从定理 9.2.1 得到.

2) 令对于 $x, y \in V$ 和 $r \in F$, $v = rx + y$,那么

$$0 = < \tau(rx + y), rx + y >$$
$$= | r |^2 < \tau(x), x > + < \tau(y), y > + r < \tau(x), y > + \bar{r} < \tau(y), x >$$
$$= r < \tau(x), y > + \bar{r} < \tau(y), x >.$$

规定 $r = 1$,得　$< \tau(x), y > + < \tau(y), x > = 0$.

规定 $r = i$,得　$< \tau(x), y > - < \tau(y), x > = 0$.

由这两个等式可得,对于 $\forall x, y \in V, < \tau(x), y > = 0$,所以由 1) 得,$\tau = 0$.

3) 考虑实内积空间 $\mathbf{R}^2$,令 $\tau \in L(V)$,由 $\tau(e_1) = e_2$ 和 $\tau(e_2) = -e_1$ 定义,知 τ 旋转了 $90°$,且对于 $\forall v, < \tau(v), v > = 0$,但 $\tau \neq 0$.

9.4 傅立叶展开

9.4.1 规范正交集

由内积,我们可以定义正交性(或垂直性)概念.

定义 9.4.1 令 V 是一个内积空间.

1) 如果 $< u, v > = 0$,称向量 $u, v \in V$ 是**正交的**,记为 $u \perp v$.

2) 如果 S 和 T 是 V 的子集,并且对于 $\forall s \in S, t \in T, s \perp t$,我们就说 S 与 T 正交,记作 $S \perp T$.

3) 子集 $S \subset T$ 的**正交补**是集合 $S^{\perp} = \{v \in V \mid v \perp S\}$.

自然有以下结论:

定理 9.4.1 令 V 是一个内积空间.

1) 对于任意子集 $S \subset V$, S 的正交补 $S^{\perp}$ 是 V 的一个子空间.

2) 对于 V 的任意子空间 S,$S\cap S^{\perp}=\{0\}$.

定义 9.4.2 在内积空间中,非空向量集 $K=\{u_i\mid i\in K\}$ 称为**正交集**,如果对于 $\forall i\neq j\in K$,$u_i\perp u_j$,且每个向量 u_i 都是单位向量,那么集合 K 是一个**规范正交集**.

因此,一个集合是规范正交的,如果对于 $\forall i, j\in K$,$<u_i, u_j>=\delta_{i,j}$,其中$\delta_{i,j}$是克罗内克 δ 函数.

给定任意非零向量 $v\in V$,给 v 乘以它的范数的倒数,就可以得到单位向量 u:

$$u=\frac{1}{\|v\|}v,$$

因此,从非零向量所成的正交集中构造一个规范正交集是很简单的.

定理 9.4.2 V 中非零向量所成的任意一个正交集都是线性无关的.

证明 令 $K=\{u_i\mid i\in K\}$ 是非零向量所成的一个正交集,设 $r_1u_1+\cdots+r_nu_n=0$,那么对 $\forall k=1, \cdots, n$,

$$0=<r_1u_1+\cdots+r_nu_n, u_k>=r_k<u_k, u_k>,$$

所以对于 $\forall k$,$r_k=0$,K 线性无关.

9.4.2 希尔伯特基和维数

定义 9.4.3 在内积空间 V 中,极大规范正交集叫做 V 的**希尔伯特基**.

可以用佐恩引理来证明任何非平凡内积空间都有一个希尔伯特基.

不要混淆向量空间的基与内积空间的希尔伯特基概念.为了避免混淆,向量空间的基,即向量的极大线性无关集,称为**哈梅尔基**.以下例子说明,通常,基的这两个概念是不同的.

例 9.4.1 令 $V=L^2$,并且令 M 是形如 $e_i=(0,\cdots,0,1,0\cdots)$ 的所有向量所成的集合,其中在 e_i 中,第 i 个坐标为 1,其余为 0.显然,M 是规范正交集,而且它是极大的.因为如果 $x=(x_n)\in L^2$ 有性质 $x\perp M$,那么对于 $\forall i$,$x_i=<x, e_i>=0$,所以 $x=0$.因此,没有一个非零向量 $x\notin M$ 与 M 正交,这说明 M 是内积空间 L^2 的一个希尔伯特基.

另一方面,M的生成向量空间是L^2 中所有含有有限个支集的序列,即,仅有有限个非零项的序列所成的子空间 S.由于 $\operatorname{span}(M)=S\neq L^2$,所以 M 不是向量空间L^2 的哈梅尔基.

我们将在后面说明一个内积空间的所有希尔伯特基都有相同的基数.

定义 9.4.4 内积空间的**希尔伯特维数**定义成这个基数.

也是为了避免混淆,V 中任意哈梅尔基的基数称为 V 的**哈梅尔维数**.

一般来说,哈梅尔维数与希尔伯特维数不同.但是,正如我们将要说明的一样,当哈

梅尔维数有限时，它们是相等的.

定义 9.4.5　令 V 是一个含有有限哈梅尔维数的内积空间. 既是 V 的哈梅尔基，又是正交集的基称为 V 的**正交哈梅尔基**；既是 V 的哈梅尔基，又是规范正交集的基称为 V 的**规范正交哈梅尔基**.（这些概念只定义在有限维向量空间上）

定理 9.4.3　令 V 是一含有有限哈梅尔维数的内积空间，则任何希尔伯特基都是一哈梅尔基. 因此，V 的希尔伯特维数与哈梅尔维数相同.

证明　令 V 有哈梅尔维数 n. 因为 V 中向量所成规范正交集是线性无关的，所以其维数不会超过 n. 特别地，一个极大规范正交集最大为 n.

如果 $\Omega = \{u_1, \cdots, u_k\}$ 是一个极大正交集，且 $k < n$，那么存在向量 $v \in V$，使得 $\Omega \cup \{v\}$ 线性无关. 如果

$$w = v + r_1 u_1 + \cdots + r_k u_k,$$

那么 $< w, u_i > = 0$，当且仅当

$$0 = < w, u_i > = < v, u_i > + r_i < u_i, u_i >,$$

或，等价地，

$$r_i = \frac{< v, u_i >}{< u_i, u_i >}.$$

用这种方式定义 r_i 后，我们得到一个与 Ω 中所有向量都正交的向量 w. 因此，$\Omega \cup \{\frac{w}{\| w \|}\}$ 是一个真包含极大规范正交集 Ω 的规范正交集. 这一矛盾说明 $k = n$，所以 Ω 是一个哈梅尔基.

如果一个内积空间含有有限希尔伯特维数，则其希尔伯特维数也等于其哈梅尔维数. 这种说法也成立. 因此，“有限维” 可以毫不含糊地应用于内积空间. 我们将用“规范正交基” 来表述规范正交哈梅尔基.

规范正交基比任意基更有优势. 为了说明这一点，设 $B = \{v_1, \cdots, v_n\}$ 是 V 的一个基，那么 $\forall v \in V$ 都有 $v = r_1 v_1 + \cdots + r_n v_n$ 形式.

但通常确定坐标 r_i 要解一个 $n \times n$ 的线性方程组.

另一方面，设 $\Omega = \{u_1, \cdots, u_n\}$ 是 V 的一个规范正交基. 任意向量 $v \in V$ 都有形式

$$v = r_1 v_1 + \cdots + r_n v_n.$$

但是现在我们有

$$< v, u_i > = < r_1 u_1 + \cdots + r_n u_n, u_i > = r_i < u_i, u_i > = r_i.$$

9.4.3　傅立叶展开

因此，在规范正交基下，找出任意向量 $v \in V$ 的坐标就非常简单了.

定理 9.4.4　令 $\Omega = \{u_1, \cdots, u_n\}$ 是 V 的规范正交基.

1) 对于 $\forall v \in V, v = <v, u_1> u_1 + \cdots + <v, u_n> u_n$. 坐标 $<v, u_i>$ 称为 v 关于 K 的**傅立叶系数**, v 右边的表达式称为 v 关于 Ω 的**傅立叶展开**.

2)（贝塞尔恒等式）　对于 $\forall v \in V$,

$$\| v \|^2 = |<v, u_1>|^2 + \cdots + |<v, u_n>|^2.$$

3)（帕塞瓦尔恒等式）　对于 $\forall v, w \in V$,

$$<v, w> = <v, u_1> \overline{<w, u_1>} + \cdots + <v, u_n> \overline{<w, u_n>}.$$

定理清楚地说明了规范正交基处理的简便性. 以下结论主要是与无限维情形下的结论做个类比.

定理 9.4.5　令 V 是一有限维内积空间. $\Omega = \{u_1, \cdots, u_k\}$ 是 V 中向量所成的一个规范正交集. 对于 $\forall v \in V$, 向量 $\tilde{v} = <v, u_1> u_1 + \cdots + <v, u_k> u_k$ 称为 v 关于 Ω 的傅立叶展开.

1)（贝塞尔不等式）　对于 $\forall v \in V, \| \tilde{v} \| \leqslant \| v \|$.

2) 集合 Ω 是 V 的一个（规范正交）哈梅尔基, 当且仅当对于 $v \in V, \tilde{v} = v$.

3) 集合 Ω 是 V 的一个（规范正交）哈梅尔基, 当且仅当对于 $\forall v \in V$, 贝塞尔恒等式成立, 即, 当且仅当对于 $\forall v \in V, \| \tilde{v} \| = \| v \|$.

4) 集合 Ω 是 V 的一个（规范正交）哈梅尔基, 当且仅当对于 $\forall v, w \in V$, 帕塞瓦尔恒等式成立, 即, 当且仅当对于 $\forall v, w \in V$,

$$<v, w> = <v, u_1> \overline{<w, u_1>} + \cdots + <v, u_k> \overline{<w, u_k>}.$$

9.5　基的正交化方法

9.5.1　射影定理

如果 S 是内积空间 V 的子空间, 那么 $S \cap S^{\perp} = \{0\}$. 这就产生了一个问题：子空间 S 的正交补是否为 S 的一个（向量空间）补, 即, 是否 $V = S \oplus S^{\perp}$.

如果 S 是 V 的一个有限维子空间, 那么答案是肯定的, 但是对于无限维子空间, S 必须有完备的拓扑性质. 我们将在后面讨论一般情况, 现在先举例说明, 通常, $V \neq S \oplus S^{\perp}$.

例 9.5.1　设 $V = L^2$, 并且令 S 是由向量 $e_i = (0, \cdots, 0, 1, 0 \cdots)$ 生成的子空间, 其中在 e_i 中, 第 i 个坐标为 1, 其余为 0. 因此, S 是 L^2 中所有含有有限个支集的序列, 即, 仅有有限个非零项的序列所成的子空间.

如果 $x = (x_n) \in S^{\perp}$, 那么对于 $\forall i, x_i = <x, e_i> = 0$, 所以 $x = 0$. 因此, $S^{\perp} =$

$\{0\}$,并且 $S \oplus S^{\perp} = S \neq L^2$.

正如下一个定理所说的那样,在有限维的情形下,正交补也是向量空间补. 通常称这个定理为射影定理,讨论到射影算子时,理由就会很显然了.(我们会在第 12 章讨论无限维情形下的射影定理).

定理 9.5.1 (射影定理) 令 S 是一内积空间的有限维子空间,则 $V = S \oplus S^{\perp}$,即,对于 $\forall v \in V$,都存在唯一向量 $s \in S$ 和 $s^{\perp} \in S^{\perp}$,使得 $v = s + s^{\perp}$.

证明 令 $\Omega = \{u_1, \cdots, u_k\}$ 是 S 的一个规范正交集. 对于 $\forall v \in V$,考虑关于 Ω 的傅立叶展开

$$\tilde{v} = \langle v, u_1 \rangle u_1 + \cdots + \langle v, u_k \rangle u_k.$$

我们可以记 $v = \tilde{v} + (v - \tilde{v})$,其中 $\tilde{v} \in S$,而且 $v - \tilde{v} \in S^{\perp}$,因为 $\langle v - \tilde{v}, u_i \rangle = \langle v, u_i \rangle - \langle \tilde{v}, u_i \rangle = 0$,所以 $V = S \oplus S^{\perp}$. 我们已知 $S \cap S^{\perp} = \{0\}$,因此 $V = S \oplus S^{\perp}$(图 9-1).

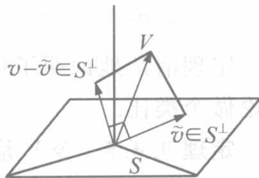


图 9-1

根据射影定理的证明,v 在 S 中的分量正好是 v 的关于 S 的任何规范正交基 Ω 的傅立叶展开.

9.5.2 正交直和

定义 9.5.1 令 V 是一内积空间,$S_1, \cdots, S_n$ 是 V 的子空间. 如果

1) $V = S_1 \oplus \cdots \oplus S_n$;

2) 对于 $i \neq j$,$S_i \perp S_j$,

那么我们就说 V 是 $S_1, \cdots, S_n$ 的**正交直和**,记为 $S = S_1 \odot \cdots \odot S_n$.

射影定理叙述了对于 V 的任意一个有限维子空间 S,$V = S \odot S^{\perp}$. 下面的简单结论非常有用:

定理 9.5.2 令 V 是一内积空间,以下论断等价:

1) $V = S \odot T$.

2) $V = S \oplus T$ 且 $T = S^{\perp}$.

3) $V = S \oplus T$ 且 $T \subset S^{\perp}$.

证明 1) ⇒2) 由 $V = S \oplus T$,$S \perp T$,这说明 $T \subset S^{\perp}$. 但是如果 $w \in S^{\perp}$,那么对于 $s \in S$,$t \in T$,$w = s + t$,所以 $0 = \langle s, w \rangle = \langle s, s \rangle + \langle s, t \rangle = \langle s, s \rangle$,这说明 $s = 0$,从而有 $w \in T$. 因此,$S^{\perp} \subset T$,从而 $S^{\perp} = T$. 所以,论断 2) 成立.

2) ⇒3) 显然成立.

3) ⇒1) 如果 3) 成立,那么 $T \subset S^{\perp}$,则 $S \perp T$,所以论断 1) 成立.

定理 9.5.3 令 V 是一内积空间.

1) 若 $\dim(V) < \infty$ 且 S 是 V 的一个子空间，那么 $\dim(S^{\perp}) = \dim(V) - \dim(S)$.

2) 如果 S 是 V 的一个有限维子空间，那么 $(S^{\perp})^{\perp} = S$.

3) 如果 S 是 V 的子集，且 $\dim(\mathrm{span}(S)) < \infty$，那么 $(S^{\perp})^{\perp} = \mathrm{span}(S)$.

证明　1) 因为 $V = S \oplus S^{\perp}$，所以 $\dim(V) = \dim(S) + \dim(S^{\perp})$.

2) 显然 $S \subset (S^{\perp})^{\perp}$. 另外，若 $v \in (S^{\perp})^{\perp}$，则由射影定理知 $v = s + s'$，其中 $s \in S$，$s' \in S^{\perp}$. 但是 $v \in (S^{\perp})^{\perp}$，所以 $0 = < v, s' > = < s', s' >$，所以 $s' = 0$，这说明 $v \in S$. 因此，$(S^{\perp})^{\perp} \subset S$，从而 $(S^{\perp})^{\perp} = S$.

9.5.3　施密特正交化方法

在内积空间 V 中给定一线性无关序列 $B = \{v_1, v_2, \cdots\}$ 后，我们在 V 中很容易就可以构造一正交序列 $\Omega = \{u_1, u_2, \cdots\}$，并且对于 $\forall k$，有性质

$$\mathrm{span}\{u_1, \cdots, u_k\} = \mathrm{span}\{v_1, \cdots, v_k\}.$$

下面的构造称为**施密特正交化方法**.

首先令 $u_1 = v_1$.

然后，我们找出形为 $u_2 = v_2 + r_1 v_1$ 的向量 u_2，使得 $< u_2, u_1 > = 0$，即，使得

$$0 = < u_2, u_1 > = < v_2 + r_1 v_1, u_1 > = < v_2, u_1 > + r_1 < u_1, u_1 >,$$

或，等价地

$$r_1 = -\frac{< v_2, u_1 >}{< u_1, u_1 >},$$

因此，用这个公式定义 r_1 后，我们看到集合 $\{u_1, u_2\}$ 是正交的，并且

$$\mathrm{span}\{u_1, u_2\} = \mathrm{span}\{v_1, v_2\}.$$

更一般地，假设 $\{u_1, \cdots, u_{k-1}\}$ 是正交的，并且

$$\mathrm{span}\{u_1, \cdots, u_{k-1}\} = \mathrm{span}\{v_1, \cdots, v_{k-1}\}.$$

想找一个形为 $u_k = v_k + r_1 u_1 + \cdots + r_{k-1} u_{k-1}$ 的向量 u_k，使得对于 $\forall i = 1, \cdots, k-1$，$< u_k, u_i > = 0$，即，

$$0 = < u_k, u_i > = < v_k + r_1 u_1 + \cdots + r_{k-1} u_{k-1}, u_i > = < v_k, u_i > + r_i < u_i, u_i >,$$

或，等价地，对于 $\forall i = 1, \cdots, j$，

$$r_i = -\frac{< v_k, u_i >}{< u_i, u_i >}.$$

由这个公式定义 r_i 后，就可以得到我们想要的向量 u_k. 下面总结一下.

定理 9.5.4(格拉姆—施密特正交化方法)　设 $B = \{v_1, v_2, \cdots\}$ 是内积空间 V 中

线性无关向量所成的一个序列. 如果我们定义

$$u_k = v_k - \sum_{i=1}^{k-1} \frac{< v_k, u_i >}{< u_i, u_i >} u_i,$$

那么序列 $\Omega = \{u_1, u_2, \cdots\}$ 是线性无关向量 v_k 所成的一个正交序列,并且对于 $\forall k = 1, 2, \cdots$ 有性质

$$\text{span}\{u_1, \cdots, u_k\} = \text{span}\{v_1, \cdots, v_k\}.$$

例 9.5.2 考虑 F 上全体多项式所成的内积空间 $F[x]$,且其内积由

$$< p(x), q(x) > = \int_{-1}^{1} p(x)q(x)\mathrm{d}x$$

定义. 在序列 $B = \{1, x, x^2, x^3, \cdots\}$ 上应用格拉姆—施密特正交化方法,得

$$u_1(x) = 1.$$

$$u_2(x) = x - \frac{\int_{-1}^{1} x\mathrm{d}x}{\int_{-1}^{1} \mathrm{d}x} \cdot 1 = x.$$

$$u_3(x) = x^2 - \frac{\int_{-1}^{1} x^2\mathrm{d}x}{\int_{-1}^{1} \mathrm{d}x} \cdot 1 - \frac{\int_{-1}^{1} x^3\mathrm{d}x}{\int_{-1}^{1} x\mathrm{d}x} \cdot x = x^2 - \frac{1}{3}.$$

$$u_4(x) = x^3 - \frac{\int_{-1}^{1} x^3\mathrm{d}x}{\int_{-1}^{1} \mathrm{d}x} \cdot 1 - \frac{\int_{-1}^{1} x^4\mathrm{d}x}{\int_{-1}^{1} x\mathrm{d}x} \cdot x - \frac{\int_{-1}^{1} x^3\left(x^2 - \frac{1}{3}\right)\mathrm{d}x}{\int_{-1}^{1}\left(x^2 - \frac{1}{3}\right)^2\mathrm{d}x} \cdot \left(x^2 - \frac{1}{3}\right)$$

$$= x^3 - \frac{3}{5}x.$$

依次类推. 这个序列中的多项式是**勒让德多项式**.

9.5.4 黎兹表示定理

如果 x 是内积空间 V 的一个向量,那么不难看出由 $\varphi_x(v) = <v, x>$ 定义的函数 $\varphi_x: V \to F$ 是 V 上一线性泛函. 以下定理表明:有限维内积空间 V 上所有线性泛函都有这种形式.

定理 9.5.5(黎兹表示定理) 令 V 是一有限维内积空间,$f \in V^*$ 是 V 上一线性泛函,那么存在唯一的向量 $\xi \in V$,使得对于 $\forall v \in V, f(v) = <v, \xi>$.

证明 如果 f 是零泛函,那么我们可以取 $\xi = 0$,所以我们假设 $f \neq 0$. 如果 $\xi \in V$,使得对于 $\forall v \in V, f(v) = <v, \xi>$,那么对于 $\forall v \in \text{Ker}(f), <v, \xi> = 0$. 因此,

我们必须在 $\mathrm{Ker}(f)^{\perp}$ 中找出一个 ξ.

如果 $\dim(V) = n$,那么 $\dim(\mathrm{Ker}(f)) = n-1$. 因此,我们可以选取单位向量 $u \in \mathrm{Ker}(f)^{\perp}$,并记为 $V = < u > \odot \mathrm{Ker}(f)$. 我们要找到一个 $r \in F$,使得对于 $\forall v \in V$, $f(v) = < v, ru >$. 特别地,对于 $v = u$,我们想得到

$$f(u) = < u, ru > = \bar{r} < u, u > = \bar{r}.$$

因此,取 $r = \overline{f(u)}$,从而 $\xi = \overline{f(u)}u$. 对于 $w \in \mathrm{Ker}(f)$,任意向量 $v \in V$ 都有形式 $v = au + bw$,所以

$$< v, \xi > = < v, \overline{f(u)}u > = f(u) < v, u > = f(u)a = f(au) = f(au+bw) = f(v).$$

由黎兹表示定理,有 $\varphi(\xi) = \xi$,可以定义映射 $\varphi: V^* \to V$,这里 ξ 是使 $f(v) = < v, ru >$ 成立的 V 中的唯一向量,即,对于 $\forall v \in V$, $\varphi(f)$ 由 $f(v) = < v, \varphi(f) >$ 定义. 因为

$$\begin{aligned} < v, \varphi(rf+sg) > &= < rf+sg > (v) = rf(v) + sg(v) \\ &= < v, \bar{r}\varphi(f) > + < v, \bar{s}\varphi(g) > \\ &= < v, \bar{r}\varphi(f) + \bar{s}\varphi(g) >, \end{aligned}$$

所以 $\varphi(rf+sg) = \bar{r}\varphi(f) + \bar{s}\varphi(g)$. 因此,$\varphi$ 是共轭线性的. 另外,φ 显然是满射的,同时也是单射的,因为 $\varphi(f) = 0$, 所以 $f = 0$,从而映射 $\varphi: V^* \to V$ 是一个共轭同构.

第 10 章　正规算子的谱理论

本章主要研究内积空间上几种特殊类型的线性算子的结构. 为了定义这些算子，我们引入另一种伴随(不同于前面的算子伴随).

本章主要讨论线性算子的伴随、正交可对角化性、自伴算子、酉算子、正规算子、正交对角化、正交射影、正交单位分解、谱定理、算子的极分解等内容.

现在，我们只在有限维的情形下定义这一伴随，而在后面再讨论无限维情形下的情况.

10.1　正交可对角化性

10.1.1　线性算子的伴随

定理 10.1.1　令 V 和 W 是 F 上有限维内积空间，且令 $\tau \in L(V,W)$，那么对于 $\forall v \in V, w \in W$，存在由条件 $<\tau(v), w> = <v, \tau^*(w)>$ 定义的唯一的函数 $\tau^*: W \rightarrow V$.

证明　对于选定的 $w \in W$，考虑由 $\theta_w(v) = <\tau(v), w>$ 定义的函数 $\theta_w: V \rightarrow F$. 容易验证 θ_w 是 V 上一个线性泛函，所以由黎兹表示定理知，存在唯一的向量 $\xi \in V$，使得对于 $\forall v \in V, \theta_w(v) = <\tau(v), w> = <v, \xi>$. 因此，如果我们规定 $\tau^*(w) = \xi$，那么对于 $\forall v \in V$，$<\tau(v), w> = <v, \tau^*(w)>$. 这就确立 τ^* 的存在和唯一性.

为了证 τ^* 是线性的，观察对于 $\forall v \in V$，

$$\begin{aligned} <v, \tau^*(rw + sw')> &= <\tau(v), rw + sw'> \\ &= \bar{r} <\tau(v), w> + \bar{s} <\tau(v), w'> \\ &= \bar{r} <v, \tau^*(w)> + \bar{s} <v, \tau^*(w')> \\ &= <v, r\tau^*(v)> + <v, s\tau^*(w')> \\ &= <v, r\tau^*(w) + s\tau^*(w')>. \end{aligned}$$

所以 $\tau^*(rw + sw') = r\tau^*(w) + s\tau^*(w')$. 因此，$\tau^* \in L(V,W)$.

定义 10.1.1　定理中定义的函数 $\tau^*: W \rightarrow V$ 在 $L(W,V)$ 中，叫做 τ 的**伴随.**

已知算子伴随 τ^* 的矩阵是映射 τ 的矩阵的转置. 对于希尔伯特空间伴随，情况有些

不同. 设 $B=\{b_1,\cdots,b_n\}$ 是 V 的有序规范正交基，$C=\{c_1,\cdots,c_m\}$ 是 W 的有序规范正交基. 如果令 $(\tau)_{B,C}=(a_{i,j})$，那么，$a_{i,j}$ 是 c_i 在 $\tau(b_j)$ 中的坐标，即 $a_{i,j}=\langle\tau(b_j),c_j\rangle$.

另一方面，如果 $(\tau^*)_{C,B}=(\alpha_{i,j})$，那么 $\alpha_{i,j}$ 是 b_i 在 $\tau^*(c_j)$ 中的坐标，即

$$\alpha_{i,j}=\langle\tau^*(c_j),b_i\rangle=\overline{\langle b_i,\tau^*(c_j)\rangle}=\overline{\langle\tau(b_i),c_j\rangle}=\overline{a}_{j,i}.$$

如果 $A=(a_{i,j})$ 是 F 上的矩阵，那么 A 的**共轭转置**是矩阵 $A^*=(\overline{a}_{i,j})^{\mathrm{T}}$.

定理 10.1.2　令 $\tau\in L(V,W)$，其中 V 和 W 是有限维内积空间，设 B 和 C 分别是 V 和 W 的有序规范正交基，那么 $(\tau^*)_{C,B}=((\tau)_{C,X})^*$，即，伴随 τ^* 的矩阵是 τ 的矩阵的共轭转置.

下面是伴随的一些基本性质：

定理 10.1.3　令 $\sigma,\tau\in L(V,W)$，其中 V 和 W 是有限维内积空间.

1) $\langle\tau^*(v),u\rangle=\langle v,\tau(u)\rangle$.

2) $(\sigma+\tau)^*=\sigma^*+\tau^*$.

3) $(r\tau)^*=\overline{r}\tau^*$.

4) $\tau^{**}=\tau^*$.

5) 如果 $V=W$，那么 $(\sigma\tau)^*=\tau^*\sigma^*$.

6) 如果 τ 是可逆的，那么 $(\tau^{-1})^*=(\tau^*)^{-1}$.

证明　3) $\langle v,(r\tau)^*u\rangle=\langle r\tau(v),u\rangle=r\langle\tau(v),u\rangle=r\langle v,\tau^*(u)\rangle=\langle v,\overline{r}\tau^*(u)\rangle$，所以 $(r\tau)^*=\overline{r}\tau^*$.

10.1.2　正交可对角化性

有限维向量空间 V 上线性算子 $\tau\in L(V)$ 是可对角化的，当且仅当 V 含有一个完全由 τ 的本征向量所成的基，或等价地，当且仅当 τ 有谱分解

$$\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k,$$

其中，$\rho_1+\cdots+\rho_k=\iota$ 是一个单位分解，λ_i 是 τ 的互异本征值，ρ_i 是映到本征空间 δ_{λ_i} 的射影. 由于

$$V=\delta_{\lambda_1}\oplus\cdots\oplus\delta_{\lambda_k},$$

所以 τ 的作用可以用以下简单形式表示出来：

$$v=v_1+\cdots+v_k\Rightarrow\tau(v)=\lambda_1v_1+\cdots+\lambda_kv_k.$$

这里对于 $\forall\, i,v_i\in\delta_{\lambda_i}$.

虽然 τ 的这种描述是简单的，但也要找出 v 的属于每个本征空间 δ_{λ_i} 的分量.

设 V 是有限维内积空间，Ω 是完全由 τ 的本征值所构成的有序规范正交基.

如果 $\Omega_i=\{u_{i,1},\cdots,u_{i,k_i}\}$ 是 Ω 的属于 λ_i 的本征值所成的子集，那么 Ω_i 是 δ_{λ_i} 的有序规范正交基，且 v 在 δ_{λ_i} 中的分量 v_i 是 v 关于 Ω_i 的易计算的傅立叶展开

$$\overline{v}_i=<v,u_{i,1}>u_{i,1}+\cdots+<v,u_{i,k_i}>u_{i,k_i}.$$

因此，τ 的作用就有了真正简单的形式：

$$v=\overline{v}_1+\cdots+\overline{v}_k\Rightarrow\tau(v)=\lambda_1\overline{v}_1+\cdots+\lambda_k\overline{v}_k.$$

定义 10.1.2 令 V 是一有限维内积空间，且 $\tau\in L(V)$，若存在 V 的一个规范正交基 Ω，使得 $(\tau)_K$ 是一个对角阵，那么我们就说 τ 是**正交可对角化**的.

显然，从讨论中可以知道 τ 是正交可对角化的，当且仅当存在 V 的一个完全由 τ 的本征向量所构成的规范正交基，即，当且仅当

$$V=\delta_{\lambda_1}\odot\cdots\odot\delta_{\lambda_k}.$$

10.2 正规算子

10.2.1 本征多项式

设 V 是 F 上一个有限维内积空间，且 $\tau\in L(V)$ 的本征多项式的根都在 F 中，即，τ 的极小多项式在 F 上分裂成线性因子的一个积，设

$$m_\tau(x)=(x-\lambda_1)^{e_1}\cdots(x-\lambda_k)^{e_k},$$

这里 λ_i 是 τ 的互异本征值(比如，在复内积空间上，任何算子都会有这样的情形)，那么，根据准素分解定理，我们把 V 记为直和 $V=V_1\oplus\cdots\oplus V_k$，这里

$$V_i=\{v\in V\mid(\tau-\lambda_i)^{e_i}(v)=0\}.$$

若 v 是 τ 的属于本征值 λ_i 的本征向量，那么 $(\tau-\lambda_i)(v)=0$，所以 $v\in V_i$. 换句话说，$\delta_{\lambda_i}\subset V_i$. 因此，$\tau$ 是正交可对角化的，当且仅当

1) 对于 $\forall i,\delta_{\lambda_i}=V_i$，且

2) 对于 $i\neq j,\delta_{\lambda_i}\perp\delta_{\lambda_j}$.

首先考虑性质 2). 这一性质等价于：对于 $i\neq j,\tau(v)=\lambda_i v$，且

$$\tau(w)=\lambda_j w\Rightarrow<v,w>=0.$$

注意到 $\lambda_i<v,w>=<\tau(v),w>=<v,\tau^*(w)>=<v,\overline{\lambda_j}w>=\lambda_j<v,w>$ (如果 $\tau^*(w)=\overline{\lambda_j}w$ 成立)，这就是说 $<v,w>=0$ (因为 $\lambda_i\neq\lambda_j$).

因此，如果对于 $\forall j,\tau$ 都有性质 $\tau(w)=\lambda_j w\Rightarrow\tau^*(w)=\overline{\lambda_j}w$，那么性质 2) 成立. 这等价于

$$(\tau-\lambda_j)(w)=0\Rightarrow(\tau^*-\overline{\lambda_j})(w)=0,$$

(或,因为$\lambda_j^*=\overline{\lambda_j}$,其中$\iota$是单位算子),所以$(\tau-\lambda_j)(w)=0\Rightarrow(\tau-\lambda_j)^*(w)=0.$

如果规定$\sigma=\tau-\lambda_j$,那么上式等价于$\sigma(w)=0\Rightarrow\sigma^*(w)=0$,从而又等价于

$$<\sigma(w),\sigma(w)>=0\Rightarrow<\sigma^*(w),\sigma^*(w)>=0.$$

如果$\sigma^*\sigma=\sigma\sigma^*$,那么上式成立.因为在这种情形下,

$$<\sigma(w),\sigma(w)>=<\sigma^*\sigma(w),w>=<\sigma\sigma^*(w),w>=<\sigma^*(w),\sigma^*(w)>.$$

但是$\sigma^*\sigma=\sigma\sigma^*$,当且仅当$\tau^*\tau=\tau\tau^*$,因此我们得出$\tau^*\tau=\tau\tau^*\Rightarrow$性质2)成立.

10.2.2　几类典型的算子

定义 10.2.1　令V是一内积空间,且令$\tau\in L(V)$,那么

如果$\tau^*=\tau$,那么τ是一个**自伴算子**,或**埃尔米特算子**.

如果τ是双射的,且$\tau^*=\tau^{-1}$,那么τ是一个**酉算子**.

如果$\tau\tau^*=\tau^*\tau$,那么τ是一个**正规算子**.

显然,自伴算子和酉算子都是正规算子.

这些定义也有矩阵的等价形,但是对于实矩阵与复矩阵,术语是不同的.如果$A=(a_{i,j})$是F上一个矩阵,那么$A^*=(\overline{a}_{i,j})^{\mathrm{T}}$是$A$的共轭转置.

定义 10.2.2　令A是一个复矩阵,那么

1）如果$A^*=A$,那么A是一个**埃尔米特矩阵**.

2）如果$A^*=-A$,那么A是一个**斜埃尔米特矩阵**.

3）如果A是可逆的,且$A^*=A^{-1}$,那么A是一个**酉矩阵**.

4）如果$AA^*=A^*A$,那么A是一个**正规矩阵**.

令A是一个实矩阵,那么$A^*=A^{\mathrm{T}}$.我们就说

5）如果$A^{\mathrm{T}}=A$,那么A是一个**对称阵**.

6）如果$A^{\mathrm{T}}=-A$,那么A是一个**斜对称阵**.

7）如果A是可逆的,且$A^{\mathrm{T}}=A^{-1}$,那么A是一个**正交阵**.

在有限维的情形下,对于V的任意有序规范正交基Ω,$(\tau^*)_\Omega=(\tau)_\Omega^*$.

所以如果τ是一个正规算子,那么

$$(\tau)_\Omega(\tau)_\Omega^*=(\tau)_\Omega(\tau^*)_\Omega=(\tau\tau^*)_\Omega=(\tau^*\tau)_\Omega=(\tau^*)_\Omega(\tau)_\Omega=(\tau)_\Omega^*(\tau)_\Omega.$$

这说明τ的矩阵$(\tau)_\Omega$是一正规矩阵.其逆命题也成立.事实上,我们可以说τ是一正规算子(自伴算子、酉算子),当且仅当关于有序规范正交基Ω的表示τ的任何矩阵都是正规矩阵(埃尔米特矩阵、酉矩阵).

10.2.3　自伴算子

定义 10.2.3　对于 $\forall v,w \in V$，$<\tau(v),w>=<v,\tau(w)>$，称算子 τ 是自伴算子.

下面是这些重要算子的一些基本性质.

定理 10.2.1　令 V 是内积空间，且令 $\sigma,\tau \in L(V)$.

1）如果 σ,τ 是自伴算子，那么 $\sigma+\tau$ 也是一个自伴算子.

2）如果 τ 是一个自伴算子，r 是一个实数，那么 $r\tau$ 是一个自伴算子.

3）如果 σ,τ 是自伴算子，那么 $\sigma\tau$ 是一个自伴算子，当且仅当 $\sigma\tau=\tau\sigma$.

4）如果 τ 是一个自伴和可逆算子，那么 τ^{-1} 也是一个自伴和可逆算子.

证明　3）注意到 $(\sigma\tau)^*=\tau^*\sigma^*$，所以 $(\sigma\tau)^*=\sigma\tau \Leftrightarrow \tau^*\sigma^*=\sigma\tau \Leftrightarrow \tau\sigma=\sigma\tau$.

定理 10.2.2　令 V 是一内积空间.

1）如果 τ 是一个自伴算子，那么对于 $\forall v\in V$，$<\tau(v),v>$ 是实数.

2）如果 V 是复内积空间，且对于 $v\in V$，$<\tau(v),v>$ 是实数，那么 τ 是一个自伴算子.

3）如果 τ 是一个自伴算子，且对于 $v\in V$，$<\tau(v),v>=0$，那么 $\tau=0$.

4）如果 τ 是一个自伴算子，那么对于 $\forall k>0$，$\tau^k(v)=0$，有 $\tau(v)=0$.

5）如果 τ 是一个自伴算子，那么 τ 的特征多项式（从而极小多项式）的复根都是实数.

6）如果 λ,μ 是自伴算子 τ 的互异本征值，那么 $\delta_\lambda \perp \delta_\mu$.

证明　1）因为 $<\tau(v),v>=<v,\tau(v)>=\overline{<\tau(v),v>}$，所以 $<\tau(v),v>$ 必为实数.

2）因为

$$
\begin{aligned}
<(\tau-\tau^*)(v),v> &= <\tau(v),v>-<\tau^*(v),v>=<\tau(v),v>-<v,\tau(v)> \\
&= <\tau(v),v>-\overline{<\tau(v),v>}=0,\text{（因为} <\tau(v),v> \text{是实数）}
\end{aligned}
$$

所以根据定理 10.2.3，$\tau-\tau^*=0$，这说明 τ 是一个自伴算子.

3）由定理10.2.3知，在复内积空间上，这是成立的，因此我们只要考虑在实内积空间上的情形，我们有

$$
\begin{aligned}
0 &= <\tau(x+y),x+y>=<\tau(x),x>+<\tau(y),y>+<\tau(x),y>+<\tau(y),x> \\
&= <\tau(x),y>+<\tau(y),x>=<\tau(x),y>+<x,\tau(y)>=<\tau(x),y>+<\tau(x),y> \\
&= 2<\tau(x),y>,
\end{aligned}
$$

所以 $\tau=0$.

4）如果对于 $\forall v \in V, \tau^k(v) = 0$，那么对于 $\forall m, \tau^{2^m}(v) = 0$，因此，

$$0 = <\tau^{2^m}(v), v> = <\tau^{2^{m-1}}\tau^{2^{m-1}}(v), v> = <\tau^{2^{m-1}}(v), \tau^{2^{m-1}}(v)>,$$

所以 $\tau^{2^{m-1}} = 0$. 重复上述步骤，最后得出 $\tau = 0$.

5）首先设 V 是复内积空间，λ 是 $C_\tau(x)$ 的一个根，那么对于 $\forall v \neq 0, \tau(v) = \lambda v$，同时

$$<\tau(v), v> = <\lambda v, v> = \lambda <v, v>;$$

$$<\tau(v), v> = <v, \tau(v)> = <v, \lambda v> = \bar{\lambda} <v, v>,$$

所以 $\lambda = \bar{\lambda}$，这说明 λ 是一个实数.

如果 V 是实向量空间，那么我们必须注意，如果 λ 是 $C_\tau(x)$ 的一个复根，并不能得出对于 $\forall 0 \neq v \in V, \tau(v) = \lambda v$. 但是我们可以如下进行下去：令 τ 由矩阵 A 表示，这里 A 是一个关于 V 的任意有序基的矩阵，那么 $C_\tau(x) = C_A(x)$. 现在，虽然 A 是实对称阵，但是我们可以把 A 看作是恰有实元的复埃尔米特阵. 同样，A 表示复空间 $\mathbf{C}^n$ 上的一个自伴线性算子，所以根据我们刚证过的内容，A 的特征多项式的所有(复) 根都是实数. 但是无论我们把 A 看作是一个实矩阵还是复矩阵，A 的特征多项式都是一样的，从而得出结论.

6）假设 $\tau(v) = \lambda v, \tau(w) = \mu w$，这里 $v, w \neq 0$，那么

$$\lambda <v, w> = <\tau(v), w> = <v, \tau(w)> = <v, \mu w> = \mu <v, w>.$$

所以 $\lambda \neq \mu$，则 $<v, w> = 0$.

当然，一个自伴算子的所有复本征值都是实数，所以 τ 的极小多项式可以分解为线性因子的积.

10.3　正交对角化

10.3.1　酉算子

现在我们讨论酉算子的基本性质。

定义 10.3.1　τ 是一个**酉算子** $\Leftrightarrow$ 对于 $\forall v, w \in V, <\tau(v), w> = <v, \tau^{-1}(w)>$.

定理 10.3.1　令 V 是一个内积空间，且令 $\sigma, \tau \in L(V)$.

1）如果 τ 是一个酉算子，那么 τ^{-1} 也是一个酉算子.

2）如果 σ, τ 是酉算子，那么 $\sigma\tau$ 也是一个酉算子.

3）τ 是一个酉算子，当且仅当它是一个满射等距.

4）如果 $\dim(V) < \infty$，那么 τ 是酉算子 $\Leftrightarrow \tau$ 可以在一个规范正交基上取一规范正交基.

5）如果 τ 是一个酉算子，那么 τ 的本征值的绝对值为 1.

证明　3）对于双射线性映射 τ，我们有

τ 是等距 $\Leftrightarrow$ 对于 $\forall v,w \in V$，$<\tau(v),\tau(w)> = <v,w>$

$\Leftrightarrow$ 对于 $\forall v,w \in V$，$<v,\tau^*\tau(w)> = <v,w>$

$\Leftrightarrow$ 对于 $\forall w \in V$，$\tau^*\tau(w) = w$

$\Leftrightarrow \tau^*\tau = \iota \Leftrightarrow \tau^* = \tau^{-1} \Leftrightarrow \tau$ 是酉算子.

4) 设 τ 是酉算子，且 $\Omega = \{u_1,\cdots,u_n\}$ 是 V 的一个规范正交基，则 $<\tau(u_i),\tau(u_j)> = <u_i,u_j> = \delta_{i,j}$，所以 $\tau(\Omega)$ 是 V 的一个规范正交基.

反之，设 K 和 $\tau(\Omega)$ 是 V 的规范正交基，那么 $<\tau(u_i),\tau(u_j)> = \delta_{i,j} = <u_i,u_j>$.

所以，如果 $v = \sum r_i u_i$，$w = \sum s_j u_j$，那么

$$<\tau(v),\tau(w)> = <\sum_i r_i\tau(u_i),\sum_j s_j\tau(u_j) = \sum_{i,j} r_i\bar{s}_j <\tau(u_i),\tau(u_j)>$$

$$= \sum_{i,j} r_i\bar{s}_j <u_i,u_j> = <\sum_i r_i u_i,\sum_j s_j u_j> = <v,w>.$$

因此，τ 是一个酉算子.

5) τ 是酉算子，且 $\tau(v) = \lambda v$，那么 $\lambda\bar{\lambda} <v,v> = <\lambda v,\lambda v> = <\tau(v),\tau(v)> = <v,v>$，所以 $|\lambda|^2 = \lambda\bar{\lambda} = 1$，则 $|\lambda| = 1$.

10.3.2 正规算子

定理 10.3.2 令 A 是一个矩阵.

1) $n \times n$ 矩阵 A 是酉矩阵 $\Leftrightarrow A$ 的列在 $\mathbf{C}^n$ 中作成一个规范正交集.

2) $n \times n$ 矩阵 A 是酉矩阵 $\Leftrightarrow A$ 的行在 $\mathbf{C}^n$ 中作成一规范正交集.

3) 如果 A 是酉矩阵，那么 $|\det(A)| = 1$.

特别地，如果 A 是正交阵，那么 $\det(A) = \pm 1$.

证明 矩阵 A 是酉矩阵 $\Leftrightarrow AA^* = I$，这与 A 的行是规范正交的这一说法等价. 同理，A 是酉矩阵 $\Leftrightarrow A^*A = I$，这与 A 的列是规范正交的这一说法等价. 有

$$AA^* = I \Rightarrow \det(A)\det(A^*) = 1 \Rightarrow \det(A)\,\overline{\det(A)} = 1.$$

我们讨论正规算子的性质，从中可以推出正规算子的定义.

定理 10.3.3 令 V 是一个内积空间，τ 是 V 上一个正规算子.

1) 对于任意多项式 $p(x) \in F[x]$，算子 $p(\tau)$ 也是一个正规算子.

2) $\tau(v) = 0 \Rightarrow \tau^*(v) = 0$.

3) 对于 $\forall k > 0$，$\tau^k(v) = 0 \Rightarrow \tau(v) = 0$.

4) 对于 $\forall \lambda \in F$，$(\tau - \lambda)^k(v) = 0 \Rightarrow (\tau - \lambda)(v) = 0$.

5) 如果 $\tau(v) = \lambda v$，那么 $\tau^*(v) = \bar{\lambda}(v)$.

6) 如果 λ,μ 是 τ 的互异本征值，那么 $\delta_\lambda \perp \delta_\mu$.

证明　3) 易知算子$\sigma=\tau\tau^*$是一个自伴算子,因为τ是一个正规算子,所以$\sigma^k(v)=(\tau^*)^k\tau^k(v)=0$,从而根据定理 10.2.2 得,$\sigma(v)=0$,即$\tau\tau^*(v)=0$.

但另一方面$0=<\tau\tau^*(v),v>=<\tau(v),\tau(v)>$,从而$\tau(v)=0$.

4) 可从论断 1) 和论断 3) 推出.

5) 设$\tau(v)=\lambda v$,其中$v\neq 0$,则$(\tau-\lambda)(v)=0$.因此,根据论断 2),$(\tau-\lambda)^*(v)=0$;同时$(\tau-\lambda)^*=\tau^*-\bar{\lambda}$,从而得出结论.

6) 假设$\tau(v)=\lambda v,\tau(w)=\mu w$,其中$v,w\neq 0$,那么$\lambda<v,w>=<\tau(v),w>=<v,\tau^*(w)>=<v,\bar{\mu}w>=\mu<v,w>$,所以$\lambda\neq\mu$,则$<v,w>=0$.

10.3.3　正交对角化

下面叙述线性代数中最完美的定理之一.

定理 10.3.4　令V是一个有限维复内积空间.

1) V上线性算子τ是正交可对角化的$\Leftrightarrow\tau$是正规算子.

2) 在V上的所有正规算子中,可以根据其本征值来描述自伴算子和酉算子.

一个正规算子是自伴算子$\Leftrightarrow$它的本征值都是实数.

一个正规算子是酉算子$\Leftrightarrow$它的本征值都有绝对值 1.

证明　1) 令τ是复内积空间上的一个正规算子.如果τ的极小多项式的素因式分解是

$$m_\tau(x)=(x-\lambda_1)^{e_1}\cdots(x-\lambda_k)^{e_k},$$

那么由准素分解定理得,$V=V_1\oplus\cdots\oplus V_k$.这里,根据定理 10.2.2,4),

$$V_i=\{v\in V\mid(\tau-\lambda_i)^{e_i}(v)=0\}=\{v\in V\mid(\tau-\lambda_i)(v)=0\}=\delta_{\lambda_i}.$$

因此$\tau|_{V_i}$的极小多项式是$x-\lambda_i$,从而对于$\forall i,e_i=1$.所以$V=\delta_{\lambda_1}\oplus\cdots\oplus\delta_{\lambda_k}$.

而且,定理 10.2.2,6) 说明V是正交直和$V=\delta_{\lambda_1}\odot\cdots\odot\delta_{\lambda_k}$.因此,由于每个本征空间都有规范正交基,我们可以构造一个τ的本征向量所成的规范正交基,所以τ是正交可对角化的.

至于逆命题,如果τ是正交可对角化的,那么存在由τ的本征向量所成的V的一个规范正交基$\Omega=\{u_1,\cdots,u_k\}$,假定$\tau(u_i)=\lambda_i u_i$,那么

$$<u_i,\tau^*(u_j)>=<\tau(u_i),u_j>=\lambda_i<u_i,u_j>=\lambda_i\delta_{i,j}=\lambda_j\delta_{i,j}=<u_i,\bar{\lambda}_j u_j>,$$

所以$\tau^*(u_j)=\bar{\lambda}_j u_j$.因此

$$\tau\tau^*(u_j)=\bar{\lambda}_j\tau(u_j)=\bar{\lambda}_j\lambda_j u_j=\lambda_j\bar{\lambda}_j u_j=\lambda_j\tau^*(u_j)=\tau^*\tau(u_j),$$

所以τ是一个正规算子.

2) 已知自伴算子是正规的,并且有实本征值.

另一方面,如果τ是正规算子,且有实本征值,那么对于任意属于本征值λ_j的本征

向量 u_j，$\tau^*(u_j)=\bar{\lambda}_j u_j=\lambda_j u_j=\tau(u_j)$. 又因为本征向量的基存在，所以 τ 是自伴算子. 因此，在有限维复内积空间上，对角阵为正规算子类作成了一典范型集（至少不计对角元的阶的情况下）. 实内积空间的情况稍有不同.

定理 10.3.5 有限维实内积空间上的线性算子 τ 是正交可对角化的，当且仅当它是一个自伴算子.

证明 设 V 是一实内积空间. 如果 τ 是一个自伴算子，那么 τ 的极小多项式在 $\mathbf{R}$ 上是分裂的. 而且，定理 10.2.2,4) 和 6)（以及运用在对称算子 $\tau-\lambda$ 上的 4)）说明 V 有一个属于 τ 的本征向量的规范正交基. 因此，τ 是正交可对角化的.

以下是逆命题的矩阵证明：如果 τ 是正交可对角化的，那么存在 V 的一个规范正交基 Ω，使得 $(\tau)_\Omega$ 是对角阵，又因为 $(\tau)_\Omega$ 是实矩阵，所以它是对称阵. 因此，$(\tau^*)_\Omega=(\tau)_\Omega^*=(\tau)_\Omega^{\mathrm{T}}=(\tau)_\Omega$. 所以 $\tau^*=\tau$.

定理 10.3.4 和 10.3.5 的矩阵等价形如下所示：

定理 10.3.6 1) 令 A 是复方阵，那么存在酉矩阵 U，使得 UAU^{-1} 是对角阵当且仅当 A 是正规矩阵.

2) 令 A 是实方阵，那么存在正交阵 T，使得 TAT^{-1} 是对角阵当且仅当 A 是对称阵.

在实域上找出酉算子的典范型的问题是：实酉算子 τ 的极小多项式 $m_\tau(x)$ 在 $\mathbf{R}$ 上可能不分裂. 如果 τ 是实酉算子，那么 $\sigma=\tau+\tau^*=\tau+\tau^{-1}$ 是自伴算子，并且有实本征值所成的一个完全集，所以我们可以分解 $V=\delta_{\lambda_1}\odot\cdots\odot\delta_{\lambda_k}$.

这里 $\delta_{\lambda_i}=\{v\in V\mid(\tau+\tau^{-1}-\lambda_i)(v)=0\}$；或乘以 τ，$\delta_{\lambda_i}=\{v\in V\mid(\tau^2-\lambda_i\tau+1)(v)=0\}$.

如果 $\lambda_i=2$，由于 τ 是正规算子，所以

$$\delta_2=\{v\in V\mid(\tau-1)^2(v)=0\}=\{v\in V\mid(\tau-1)(v)=0\}.$$

如果 $\lambda_i=-2$，那么

$$\delta_{-2}=\{v\in V\mid(\tau+1)^2(v)=0\}=\{v\in V\mid(\tau+1)(v)=0\}.$$

因此，在本征空间 δ_2 和 δ_{-2} 上（如果它们确实存在），算子 τ 正好分别乘以了 1 或 -1.

对于 $\lambda_i\neq\pm2$，我们可以如下分解每个 δ_{λ_i}：取 $v\in\delta_{\lambda_i}$，考虑 $\mathrm{span}\{v,\tau(v)\}$，$\delta_{\lambda_i}$ 的这个子空间是不变的，因为 $\tau(\tau(v))=\tau^2(v)=\lambda_i\tau(v)-v$. 因此，我们可以记为

$$\delta_{\lambda_i}=\mathrm{span}\{v,\tau(v)\}\odot\mathrm{span}\{v,\tau(v)\}^\perp.$$

这样一直继续下去，我们就可以把每个 δ_{λ_i} 记为二维空间的一个正交直和，其中 τ 是一个实酉算子，从而

$$V=\delta_2\odot\delta_{-2}\odot W_1\odot\cdots\odot W_m.$$

这里 $\dim(W_i)=2$，并且每个直和项在 τ 之下都是不变的.

因此,我们只要确定二维空间 W 上实酉算子 τ 的矩阵. τ 的关于 W 的任意规范正交基的矩阵都是正交的,所以如果

$$(\tau)=\begin{pmatrix}a & b\\ c & d\end{pmatrix},$$

那么 $a^2+b^2=1, c^2+d^2=1, ac+bd=0$. 又因为 $\det(\tau)$ 是极小多项式

$$m_\tau(x)=x^2-\lambda_i\tau+1=0$$

的常数项,所以 $\det(\tau)=1$,即 $ad-bc=1$. 解这些方程得 $d=a, c=-b$,所以

$$(\tau)=\begin{pmatrix}a & b\\ -b & a\end{pmatrix}.$$

因为 (a,b) 是 $\mathbf{R}^2$ 中的单位向量,所以对于任意实数 θ, $(a,b)=(\cos\theta,\sin\theta)$,从而

$$(\tau)=\begin{pmatrix}\cos\theta & \sin\theta\\ -\sin\theta & \cos\theta\end{pmatrix}.$$

因此,我们就有了以下

定理 10.3.7　令 τ 是有限维实内积空间 V 上的一个酉算子,那么存在 V 的一规范正交基,使得 τ 的矩阵有分块形式

$$\begin{pmatrix}1 &&&&&&&&\\ &\ddots &&&&&&&\\ &&1&&&&&&\\ &&&-1&&&&&\\ &&&&\ddots&&&&\\ &&&&&-1&&&\\ &&&&&&\begin{bmatrix}\cos\theta_1 & \sin\theta_1\\ -\sin\theta_1 & \cos\theta_1\end{bmatrix}&&\\ &&&&&&&\ddots&\\ &&&&&&&&\begin{bmatrix}\cos\theta_k & \sin\theta_k\\ -\sin\theta_k & \cos\theta_k\end{bmatrix}\end{pmatrix}$$

10.4　线性算子的正交分解

10.4.1　正交射影

现在我们从射影算子方面来描述正交可对角化性的特征.

定义 10.4.1　令 $V = S \odot S^{\perp}$. S 上沿 $S^{\perp}$ 的射影映射 ρ: $V \to S$ 叫做映到 S 上的**正交射影**. 换句话说,射影映射 ρ 是一个正交射影,如果 $V = \mathrm{Im}(\rho) \odot \mathrm{Ker}(\rho)$.

因此,正交射影恰是一种特殊的射影算子,其中 $\mathrm{Ker}(\rho) = \mathrm{Im}(\rho)^{\perp}$. 要特别注意的是,不要混淆术语"正交射影"与概念"互相正交的射影",即 $\rho\sigma = \sigma\rho = 0$.

算子 ρ 是一个射影算子 $\Leftrightarrow \rho$ 是一个幂等算子.

定理 10.4.1　线性算子 $\rho \in L(V)$ 是一个正交射影 $\Leftrightarrow \rho$ 既是幂等又是自伴算子.

证明　设 ρ 既是幂等算子又是自伴算子,那么 ρ 是 $\mathrm{Im}(\rho)$ 上沿 $\mathrm{Ker}(\rho)$ 的射影, $V = \mathrm{Im}(\rho) \oplus \mathrm{Ker}(\rho)$. 而且,如果 $x \in \mathrm{Ker}(\rho)$,那么 $< \rho(v), x > = < v, \rho(x) > = 0$. 所以 $\mathrm{Im}(\rho) \perp \mathrm{Ker}(\rho)$. 因此, $V = \mathrm{Im}(\rho) \odot \mathrm{Ker}(\rho)$,这说明 ρ 是一个正交射影.

假设 ρ 是一个正交射影,那么 ρ 就是一个幂等算子,我们只要证它是一个自伴算子. 因为 ρ 是一个正交射影,所以 $V = \mathrm{Im}(\rho) \odot \mathrm{Ker}(\rho)$. 但是如果 $v \in \mathrm{Im}(\rho)$,那么 $v = \rho(w)$,从而

$$\rho(v) = \rho(\rho(w)) = \rho^2(w) = \rho(w) = v.$$

因此, $\mathrm{Im}(\rho)$ 中所有非零向量都是属于本征值 1 的本征向量. 而且,如果 $x \in \mathrm{Ker}(\rho)$,那么 $\rho(x) = 0 = 0x$. 所以, $\mathrm{Ker}(\rho)$ 中所有非零向量都是属于本征值 0 的本征向量. 因此,我们可以找出一个完全由 ρ 的本征向量所成的 V 的规范正交基,这说明 ρ 是一个正规算子.

最后,因为 ρ 的本征值是实数,所以 ρ 必是一个自伴算子. 对于正交射影 ρ,我们有

$$< v, \rho(v) > = < v, \rho^2(v) > = < \rho(v), \rho(v) >.$$

10.4.2　正交射影的性质

定理 10.4.2　线性算子 $\rho \in L(V)$ 是一个正交射影 $\Leftrightarrow \rho$ 是一个幂等算子,并且对于 $\forall v \in V$, $\| \rho(v) \| \leqslant \| v \|$.

证明　充分性　假设 ρ 是一个幂等算子,且 $\| \rho(v) \| \leqslant \| v \|$.

我们要证 $V = \mathrm{Ker}(\rho) \odot \mathrm{Im}(\rho)$,根据定理 10.4.2,只要证 $\mathrm{Im}(\rho) \subset \mathrm{Ker}(\rho)^{\perp}$. 这个问题的关键是 $V = \mathrm{Ker}(\rho) \odot \mathrm{Ker}(\rho)^{\perp}$,由射影定理,对于 $\dim(\mathrm{Ker}(\rho)) < \infty$,这是成立的. 在这一假设之下继续下去,从而对于 $\forall w \in \mathrm{Im}(\rho)$, $w = x + y$,其中 $x \in \mathrm{Ker}(\rho)$, $y \in \mathrm{Ker}(\rho)^{\perp}$,由于 ρ 是一个幂等算子, $w = \rho(w) = \rho(x) + \rho(y) = \rho(y)$,所以

$$\| x \|^2 + \| y \|^2 = \| w \|^2 = \| \rho(y) \|^2 \leqslant \| y \|,$$

则 $\| x \| = 0$,从而 $x = 0$. 所以 $w = y \in \mathrm{Ker}(\rho)^{\perp}$,从而 $\mathrm{Im}(\rho) \subset \mathrm{Ker}(\rho)^{\perp}$.

下面三个定理给出了正交射影的一些其他性质.

定理 10.4.3　1) 如果 ρ 和 σ 都是正交射影,那么由 $\rho\sigma = 0$ 可得 $\sigma\rho = 0$.

2) 两正交射影 ρ 和 σ 互相正交,当且仅当 $\mathrm{Im}(\rho)\perp\mathrm{Im}(\sigma)$.

定理 10.4.4　令 V 是域上特征 $\neq 2$ 的一个向量空间.

1) 令 ρ 和 σ 是正交射影,则 $\rho+\sigma$ 是一正交射影当且仅当 $\rho\perp\sigma$,在此情形下 $\rho+\sigma$ 是 $\mathrm{Im}(\rho)\odot\mathrm{Im}(\sigma)$ 上沿 $\mathrm{Ker}(\rho)\cap\mathrm{Ker}(\sigma)$ 的射影.

2) 令 $\rho_1,\cdots,\rho_k$ 是正交射影,那么 $\rho=\rho_1+\cdots+\rho_k$ 是一个正交射影,当且仅当对于 $\forall i\neq j,\rho_i\perp\rho_j$.

3) 令 ρ 和 σ 都是正交射影,那么 $\rho-\sigma$ 是一个正交射影,当且仅当 $\rho\sigma=\sigma\rho=\sigma$. 在这一情形下,$\rho-\sigma$ 是 $\mathrm{Im}(\rho)\cap\mathrm{Ker}(\sigma)$ 上沿 $\mathrm{Ker}(\rho)\odot\mathrm{Im}(\sigma)$ 的射影.

4) 令 ρ 和 σ 是正交射影,若 $\rho\sigma=\sigma\rho$,则 $\rho\sigma$ 是一个正交射影. 在此情形下,$\rho\sigma$ 是 $\mathrm{Im}(\rho)\cap\mathrm{Im}(\sigma)$ 上沿 $\mathrm{Ker}(\rho)\odot\mathrm{Ker}(\sigma)$ 的射影.

证明　2) 如果 ρ_i 是一个正交射影,且对于 $\forall i\neq j,\rho_i\perp\rho_j$,那么对于 $\forall i\neq j,\rho_i\rho_j=0$,所以不难验证 $\rho^2=\rho,\rho^*=\rho$. 因此,ρ 是一个正交射影.

反之,假设 ρ 是一个正交射影,$x\in\mathrm{Im}(\rho_i)$,那么 $\rho_i(x)=x$,从而

$$\begin{aligned}\|x\|^2&\geqslant\|\rho(x)\|^2=<\rho(x),\rho(x)>=<\rho(x),x>\\&=\sum_j<\rho_j(x),x>=\sum_j\|\rho_j(x)\|^2\\&\geqslant\|\rho_i(x)\|^2=\|x\|^2,\end{aligned}$$

则对于 $j\neq i,\rho_j(x)=0$. 换句话说,$\mathrm{Im}(\rho_i)\subset\mathrm{Ker}(\rho_j)=\mathrm{Im}(\rho_j)^\perp$. 因此,对于 $\forall v,w\in V,0=<\rho_j(v),\rho_i(w)>=<\rho_i\rho_j(v),w>$. 这说明 $\rho_i\rho_j=0$,即 $\rho_i\perp\rho_j$.

定理 10.4.5　对于正交投射影 ρ 和 σ,以下论断等价:

1) 对于 $\forall v\in V,<(\rho-\sigma)(v),v>\geqslant 0$.

2) 对于 $\forall v\in V,\|\sigma(v)\|\leqslant\|\rho(v)\|$.

3) $\mathrm{Im}(\sigma)\subset\mathrm{Im}(\rho)$.

4) $\rho\sigma=\sigma$.

5) $\sigma\rho=\sigma$.

如果这些条件中的任意一个成立,我们就说 σ 小于或等于 ρ,记为 $\sigma\leqslant\rho$.

证明　1)⇒2)　由 1) 成立,有

$$\begin{aligned}0&\leqslant<(\rho-\sigma)(v),v>=<(\rho(v),v>-<\sigma(v),v>\\&=<(\rho(v),\rho(v)>-<\sigma(v),\sigma(v)>\\&=\|\rho(v)\|^2-\|\sigma(v)\|^2.\end{aligned}$$

2)⇒3)　由 2) 成立,对于 $\forall v\in\mathrm{Im}(\sigma),v=x+y$,这里 $x\in\mathrm{Im}(\rho)\perp y\in\mathrm{Ker}(\rho)$,

那么

$$\| x \|^2 + \| y \|^2 = \| v \|^2 = \| \sigma(v) \|^2 \leqslant \| \rho(v) \|^2 = \| x \|^2.$$

所以 $y = 0$,即 $v \in \mathrm{Im}(\rho)$.

3)⇒4) 由 3) 成立,因为对于 $\forall v \in V, \sigma(v) \in \mathrm{Im}(\sigma) \subset \mathrm{Im}(\rho)$,所以 $\rho(\sigma(v)) = \sigma(v)$. 从而 $\rho\sigma = \sigma$.

4)⇒5) 由 4) 成立,有 $\sigma\rho = \sigma^* \rho^* = (\rho\sigma)^* = \sigma^* = \sigma$.

5)⇒1) 由 3),假设 5) 成立,4) 也成立,所以 $\rho - \sigma$ 是一个正交投射影,从而

$$< (\rho - \sigma)(v), v > = < (\rho - \sigma)(v), (\rho - \sigma)(v) > \geqslant 0.$$

10.4.3 正交单位分解

定义 10.4.2 如果 $\rho_1, \cdots, \rho_k$ 是使得 $\rho_1 + \cdots + \rho_k = \iota$ 为一个单位分解的正交射影,那么我们就称 $\rho_1 + \cdots + \rho_k = \iota$ 为**正交单位分解**.

定理 10.4.6 1) 如果 $\rho_1 + \cdots + \rho_k = \iota$ 是一个正交单位分解,那么

$$V = \mathrm{Im}(\rho_1) \odot \cdots \odot \mathrm{Im}(\rho_k).$$

2) 反之,若 $V = S_1 \odot \cdots \odot S_k$, ρ_i 是 S_i 上沿 $S_1 \odot \cdots \odot \hat{S}_i \cdots \odot S_k$ 的射影,其中ˆ指的是其对应项从直和中去掉,那么 $\rho_1 + \cdots + \rho_k = \iota$ 是正交单位分解.

证明 1) 设 $\rho_1 + \cdots + \rho_k = \iota$ 是一个正交单位分解,则

$$V = \mathrm{Im}(\rho_1) \oplus \cdots \oplus \mathrm{Im}(\rho_k).$$

但是,由于 ρ_i 是正交射影,所以它们是自伴的,因此对于 $i \neq j$, $< \rho_i(v), \rho_j(w) > = < v, \rho_i\rho_j(w) > = 0$. 这说明 $V = \mathrm{Im}(\rho_1) \odot \cdots \odot \mathrm{Im}(\rho_k)$.

逆命题,因 $\rho_1 + \cdots + \rho_k = \iota$ 是一个单位分解,因为

$$\mathrm{Im}(\rho_i) = S_i \perp (S_1 \odot \cdots \odot \widehat{S}_i \cdots \odot S_k) = \mathrm{Ker}(\rho_i),$$

所以每一个 ρ_i 都是一个正交射影.

10.5 线性算子的谱理论

10.5.1 正规算子的谱定理

现在,我们可以在有限维复内积空间上描述正交可对角化算子了.

定理 10.5.1(正规算子的谱定理) 令 $\tau \in L(V)$,其中 V 是有限维复内积空间,以下论断等价:

1）τ 是正交可对角化算子，即 $V=\delta_{\lambda_1}\odot\cdots\odot\delta_{\lambda_k}$.

2）τ 是正规算子，即 $\tau\tau^*=\tau^*\tau$.

3）τ 有正交谱分解 $\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k$，其中 $\lambda_i\in\mathbf{C}$，$\rho_1+\cdots+\rho_k=\iota$ 是一个正交单位分解.

而且，如果 τ 有形式 $\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k$，其中 λ_i 互异，ρ_i 不为零，那么 λ_i 是 τ 的本征值，$\mathrm{Im}(\rho_i)$ 是 τ 的属于 λ_i 的本征空间.

证明　由定理 10.3.4，论断 1）与论断 2）等价.

3）设 τ 是正交可对角化的. 对于任一单位分解，$\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k$ 成立，因为 $V=\delta_{\lambda_1}\odot\cdots\odot\delta_{\lambda_k}$，所以这是一个正交分解. 从而论断 3）成立.

反之，如果 $\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k$ 成立，那么 $V=\mathrm{Im}(\rho_1)\odot\cdots\odot\mathrm{Im}(\rho_k)$. 则 $\mathrm{Im}(\rho_i)=\delta_{\lambda_i}$，所以 τ 是正交可对角化的.

在有限维实内积空间中，我们有以下

定理 10.5.2(自伴算子的谱定理)　令 $\tau\in L(V)$，V 是有限维复内积空间，以下论断等价.

1）τ 是正交可对角化算子，即 $V=\delta_{\lambda_1}\odot\cdots\odot\delta_{\lambda_k}$.

2）τ 是自伴算子，即 $\tau^*=\tau$.

3）τ 有正交谱分解 $\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k$，其中 $\lambda_i\in\mathbf{R}$，$\rho_1+\cdots+\rho_k=\iota$ 是一个正交单位分解.

而且，如果 τ 有形式 $\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k$，其中 λ_i 互异，ρ_i 不为零，那么 λ_i 是 τ 的本征值，$\mathrm{Im}(\rho_i)$ 是 τ 的属于本征值 λ_i 的本征空间.

10.5.2　谱定理的应用

如果 V 是 F 上一向量空间，$\tau\in L(V)$，$p(x)$ 是 F 上一个多项式，那么算子 $p(\tau)\in L(V)$ 是唯一确定的. 现在，设 V 是一个有限维内积空间，τ 有谱分解

$$\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k,$$

那么对于 $m\geqslant 1$，$\rho_i^m=\rho_i$，且对于 $i\neq j$，$\rho_i\rho_j=0$. 因此，

$$\tau^n=(\lambda_1\rho_1+\cdots+\lambda_k\rho_k)^n=\lambda_1^n\rho_1+\cdots+\lambda_k{}^n\rho_k.$$

更一般地，对于 F 上任意多项式 $p(x)$，$p(\tau)=p(\lambda_1)\rho_1+\cdots+p(\lambda_k)\rho_k$.

事实上，对于任意函数

$$f:\{\lambda_1,\cdots,\lambda_k\}\to F,$$

通过定义 $f(\tau)=f(\lambda_1)\rho_1+\cdots+f(\lambda_k)\rho_k$，我们可以进一步扩充上式. 因此，我们可以定

义$\sqrt{\tau},\tau^{-1},e^{\tau}$,等等.

因为τ的谱分解是一个有限和,所以在运用函数,而不是多项式时,实际上我们什么也没有得到(除了方便外).

设$f:\{\lambda_1,\cdots,\lambda_k\}\rightarrow F$是任意一个函数,令$f(\lambda_i)=\alpha_i$,则我们可以找到一多项式$p(x)$,使得对于$i=1,\cdots,k,p(\lambda_i)=\alpha_i$.因此,

$$f(\tau)=f(\lambda_1)\rho_1+\cdots+f(\lambda_k)\rho_k=p(\lambda_1)\rho_1+\cdots+p(\lambda_k)\rho_k=p(\tau).$$

对算子τ的函数性质的研究称为τ的函数演算.

如果V是复内积空间($F=\mathbf{C}$),τ是正规算子,那么$f(\tau)$是一个正规算子,其本征值为$f(\lambda_i)$.同理,如果V是一个实内积空间($F=\mathbf{R}$),τ是一个自伴算子,那么$f(\tau)$是一个自伴算子,其本征值为$f(\lambda_i)$.我们考虑一下这种构造的几个特例.

对于$\forall j=1,\cdots,k$,如果$p_j(x)$是多项式,使得$p_j(\lambda_j)=1,p_j(\lambda_i)=0$,对于$i\neq j$,有$p_j(\tau)=\rho_j$.从而,每个射影$\rho_j$都是$\tau$的一个多项式函数.

若τ是可逆的,则对于$\forall i,\lambda_i\neq 0$,可令$f(x)=x^{-1}$,得

$$\tau^{-1}=\lambda_1^{-1}\rho_1+\cdots+\lambda_k^{-1}\rho_k.$$

若$f(\lambda_i)=\overline{\lambda}_i$,$\tau$是一个正规算子,则每一$\rho_i$都是自伴的,所以

$$f(\tau)=\overline{\lambda}_1\rho_1+\cdots+\overline{\lambda}_k\rho_k=\tau^*.$$

函数演算可以运用到算子的交换性性质的研究上.下面是两个简单的例子.

定理 10.5.3 令τ有谱分解$\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k$,那么算子$\sigma$与$\tau$交换$\Leftrightarrow\sigma$与每一个$\rho_i$可交换.

证明 如果σ与每个ρ_i交换,那么显然σ与τ交换.至于逆命题,观察到ρ_i是τ中的一个多项式,由于σ与τ交换,所以它与τ中任意多项式交换.

定理 10.5.4 令V是一有限维复内积空间,$\tau,\sigma\in L(V)$是正规算子,那么τ和σ可交换$\Leftrightarrow$它们有形式$\tau=p(\theta),\sigma=q(\theta)$,这里$p$和$q$是多项式,$\theta=r(\tau,\sigma)$是$\tau$和$\sigma$的一个多项式.

证明 如果τ和σ是θ中的多项式,那么它们显然交换.至于逆命题,设$\tau\sigma=\sigma\tau$,令

$$\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k;\sigma=\mu_1\upsilon_1+\cdots+\mu_m\upsilon_m$$

是τ和σ的正交谱分解,那么,$\rho_i\upsilon_j=\upsilon_j\rho_i$.选取任意多项式$r(x,y)$,有性质$\alpha_{i,j}=r(\lambda_i,\mu_j)$是互异的.因为每一个$\rho_i$和$\upsilon_j$都是自伴的,所以可以规定$\theta=r(\tau,\sigma)$,从而推出

$$\theta=r(\tau,\sigma)=\sum_{i,j}\alpha_{i,j}\rho_i\upsilon_j.$$

同样,选取$p(x)$和$q(x)$,使得对于$\forall j,p(\alpha_{i,j})=\lambda_i$,对于$\forall i,q(\alpha_{i,j})=\mu_j$.那么

$$p(\theta)=\sum_{i,j}p(\alpha_{i,j})\rho_i\upsilon_j=\sum_{i,j}\lambda_i\rho_i\upsilon_j=(\sum_i\lambda_i\rho_i)(\sum_j\upsilon_j)=\sum_i\lambda_i\rho_i=\tau.$$

同理，$q(\theta)=\sigma$.

10.5.3　正算子

重要的函数演算例子之一是当 $f(x)=\sqrt{x}$ 时. 首先需要一些定义.

定义 10.5.1　自伴线性算子 $\tau\in L(V)$ 是**非负的**，若对 $\forall v\in V$，$<\tau(v),v>\geqslant 0$；$\tau\in L(V)$ 是**正的**，若 τ 是非负的，且对 $v\neq 0$，$<\tau(v),v>>0$.

定理 10.5.5　有限维内积空间上自伴算子 τ 是

1) 非负的，当且仅当 τ 的本征值都是非负的.

2) 正的，当且仅当 τ 的本征值都是正的.

证明　如果 $<\tau(v),v>\geqslant 0$，$\tau(v)=\lambda v$，那么 $0\leqslant<\tau(v),v>=\lambda<v,v>$，所以 $\lambda\geqslant 0$.

反之，如果 τ 的本征值都是非负的，那么

$$\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k,\lambda_i\geqslant 0.$$

因为 $\iota=\rho_1+\cdots+\rho_k$，

$$<\tau(v),v>=\sum_{i,j}\lambda_i<\rho_i(v),\rho_j(v)>=\sum_i\lambda_i\|\rho_i(v)\|^2\geqslant 0,$$

所以 τ 是非负的.

如果 τ 是一个非负算子，且有谱分解

$$\tau=\lambda_1\rho_1+\cdots+\lambda_k\rho_k,\lambda_i\geqslant 0,$$

那么我们可以取 τ 的非负平方根

$$\sqrt{\tau}=\sqrt{\lambda_1}\rho_1+\cdots+\sqrt{\lambda_k}\rho_k,$$

其中 $\sqrt{\lambda_i}$ 是 λ_i 的非负平方根. 显然，$(\sqrt{\tau})^2=\tau$. 而且不难看出 $\sqrt{\tau}$ 是平方为 τ 的唯一的非负算子. 换句话说，每个非负算子都有唯一的非负平方根. 反之，如果 τ 有一个非负平方根，即，如果对于任意非负算子 σ，$\tau=\sigma^2$，那么 τ 是非负的. 因此，算子 τ 是非负的，当且仅当 τ 有一个非负平方根.

下面是平方根的一个应用：

定理 10.5.6　若 τ 和 σ 是非负算子，$\tau\sigma=\sigma\tau$，则 $\tau\sigma$ 是非负的.

证明　因为 τ 是一个非负算子，所以 τ 有一个非负平方根 $\sqrt{\tau}$，且 $\sqrt{\tau}$ 是 τ 中的一个多项式，同样 σ 也有非负平方根 $\sqrt{\sigma}$. 由于 τ 和 σ 交换，所以 $\sqrt{\tau}$ 和 $\sqrt{\sigma}$ 也交换，从而，

$$(\sqrt{\tau}\sqrt{\sigma})^2 = (\sqrt{\tau})^2(\sqrt{\sigma})^2 = \tau\sigma.$$

由于$\sqrt{\tau}$和$\sqrt{\sigma}$是自伴且交换的，所以它们的积是自伴的，从而$\tau\sigma$是非负的.

10.5.4　线性算子的极分解

任意非零复数z都可以记为极形式$z = re^{i\theta}$，这里r是一正数，θ是一实数. 在有限维复内积空间上，这也适用于任意非零线性算子τ.

定理 10.5.7　令τ是有限维复内积空间V上一个非零线性算子，那么存在一个酉算子υ和唯一的正算子ρ，使得$\tau = \upsilon\rho$. 而且，如果τ是可逆的，那么υ就是唯一的.

证明　设$\tau = \upsilon\rho$，那么$\tau^* = (\upsilon\rho)^* = \rho^*\upsilon^* = \rho\upsilon^{-1}$. 所以$\tau^*\tau = \rho\upsilon^{-1}\upsilon\rho = \rho^2$. 同样，如果$v \in V$，那么$\tau(v) = \upsilon\rho(v)$. 这些等式提示了我们怎样定义$\rho$和$\upsilon$.

我们定义ρ为非负算子$\tau^*\tau$的唯一的非负平方根，那么

$$\|\rho(v)\|^2 = \langle\rho(v),\rho(v)\rangle = \langle\rho^2(v),v\rangle = \langle\tau^*\tau(v),v\rangle = \|\tau(v)\|^2 \quad (10.1)$$

对于$\forall v \in V$，由$\upsilon(\rho(v)) = \tau(v)$，我们在象$\mathrm{Im}(\rho)$上定义$\upsilon$. 注意到上式，

$$\rho(v) = \rho(w) \Rightarrow \rho(v-w) = 0 \Rightarrow \|\rho(v-w)\| = 0 \Rightarrow \|\tau(v-w)\| = 0 \Rightarrow \tau(v) = \tau(w).$$

而且，υ(在它的定义域$\mathrm{Im}(\rho)$上)是一个等距，又由(10.1)式得，

$$\|\upsilon(\rho(v))\| = \|\tau(v)\| = \|\rho(v)\|.$$

因为υ: $\mathrm{Im}(\rho) \to \mathrm{Im}(\upsilon)$是单射的，所以

$$\dim(\mathrm{Im}(\rho)) = \dim(\mathrm{Im}(\upsilon)).$$

从而

$$\dim(\mathrm{Im}(\rho)^{\perp}) = \dim(\mathrm{Im}(\upsilon)^{\perp}).$$

这样就可以把υ扩充到V的一个酉映射υ上. 这样，等式$\upsilon(\rho(v)) = \tau(v)$成立，就说明$\tau = \upsilon\rho$.

至于唯一性，假设$\tau = \upsilon\rho = \upsilon'\rho'$，那么

$$\tau^*\tau = \rho^*\upsilon^*\upsilon\rho = \rho^2, \tau^*\tau = \rho'\upsilon'^*\upsilon'\rho' = (\rho')^2.$$

所以$\rho^2 = (\rho')^2$，因为ρ^2有唯一的非负平方根，所以我们推出$\rho = \rho'$. 因此ρ是唯一的. 最后，如果τ是可逆的，那么(10.1)式说明ρ也是可逆的. 因此，ρ是一个双射，所以$\upsilon(\rho(v)) = \tau(v)$，唯一确定$\upsilon$.

把前一个定理运用到映射τ^*上，$\tau = (\tau^*)^* = (\upsilon\rho)^* = \rho\upsilon^{-1} = \rho\mu$. 这样就有了以下

推论 10.5.8　令τ是有限维复内积空间上一个非零线性算子，那么存在自伴算子σ

和唯一的正算子 ρ，使得 τ 有**极分解** $\tau = \rho e^{i\sigma}$.

我们可以运用极分解来描述正规算子的特征.

定理 10.5.9　令 $\tau = \rho e^{i\sigma}$ 是非零线性算子 τ 的一个极分解，那么 τ 是一个正规算子，当且仅当 $\rho\sigma = \sigma\rho$.

证明　因为 $\tau\tau^* = \rho e^{i\sigma} e^{-i\sigma}\rho$；$\tau^*\tau = e^{-i\sigma}\rho\rho e^{i\sigma} = e^{-i\sigma}\rho^2 e^{i\sigma}$. 由于 τ 是正规算子，当且仅当 $e^{-i\sigma}\rho^2 e^{i\sigma} = \rho^2$. 或等价地 $\rho^2 e^{i\sigma} = e^{i\sigma}\rho^2$. 现在，$\rho$ 是 ρ^2 的一个函数，σ 是 $e^{i\sigma}$ 的一个函数，所以 $\rho^2 e^{i\sigma} = e^{i\sigma}\rho^2 \Leftrightarrow \rho\sigma = \sigma\rho$.

第11章 度量线性空间

本章主要讨论双线性型的矩阵、二次型、线性泛函、正交直和、商空间、辛几何—双曲平面、正交几何的结构、维特消去定理、维特扩张定理、极大双曲子空间等内容.

11.1 双线性型的矩阵

11.1.1 对称型、斜对称型与交错型

本章中,我们将研究有双线性型的任意域上的向量空间. 正如我们将要看到的,这种向量空间的研究有着很强的几何味道.

除非特别声明,所有向量空间都假设为有限维的,符号 F 表示任意域,F_q 表示含有 q 个元素的有限域.

定义 11.1.1 令 V 是 F 上的向量空间. 映射 $<,>: V\times V\to F$ 叫做一个**双线性型**,如果它是每个坐标的线性函数,即,如果

$$<\alpha x+\beta y, z> = \alpha <x,z> + \beta <y,z>;$$

$$<z, \alpha x+\beta y> = \alpha <z,x> + \beta <z,y>.$$

一个双线性型是

1) **对称的**,如果对于 $\forall x,y\in V$, $<x,y>=<y,x>$.

2) **斜对称的**,如果对于 $\forall x,y\in V$, $<x,y>=-<y,x>$.

3) **交错的**,如果对于 $\forall x\in V$, $<x,x>=0$.

定义 11.1.2 对称型、斜对称型或交错型的双线性型称为**内积**,$(V,<,>)$ 叫做**度量线性空间**,这里 V 是一个向量空间,$<,>$ 是 V 上的一个内积.

按照上述定义,前面所讨论的实内积就是内积,并且有附加性质 —— 是正定的. 而本章所讨论的复内积是半双线性的,按照上述定义,它并不是内积. 还要注意的是,不要把度量线性空间与我们下一章将要研究的距离空间混为一谈.

提出内积的定义后,我们将和前面一样,运用"令 V 是一个度量线性空间".

定义 11.1.3 令 V 是域 F 上一度量线性空间. 如果 $<,>$ 是对称的,那么 V 叫做 F 上的**正交几何**,如果 $<,>$ 是交错的,那么 V 叫做 F 上的**辛几何**.

因此,实内积空间是正交几何,但复内积空间不是正交几何.

并不是所有度量线性空间都和实内积空间一样表现得那么完美,这样我们就有必要引入新的术语,以此来适应其不同的表现形式.下面是一个例子.

如果对于 $\forall v \in V$, $<x,v>=0 \Rightarrow x=0$,称度量线性空间 V 是**非奇异的**.

例 11.1.1　闵可夫斯基空间 $\mathbf{M}_4$ 的内积由 $<e_1,e_1>=<e_2,e_2>=<e_3,e_3>=1$, $<e_4,e_4>=-1$,且对于 $i \neq j$, $<e_i,e_j>=0$ 定义的四维非奇异实正交几何 $\mathbf{R}^4$,这里 $e_1,\cdots,e_4$ 是 $\mathbf{R}^4$ 的标准基.

虽然对称型、斜对称型和交错型的概念不是各自独立的,但它们之间的关系依赖于基域 F 的特征.

定理 11.1.1　令 V 是域 F 上一向量空间.

1) 如果 $\mathrm{char}(F)=2$,那么 V 上的双线性型是斜对称的,当且仅当它是对称的.而且,交错双线性型是对称(和斜对称)的.

2) 如果 $\mathrm{char}(F)\neq 2$,那么 V 上的双线性型是斜对称的 $\Leftrightarrow$ 它是交错的.

证明　首先,我们观察到对于任意域,如果 $<,>$ 是对称的,那么

$$0=<x+y,x+y>=<x,x>+<x,y>+<y,x>+<y,y>$$
$$=<x,y>+<y,x>.$$

所以,$<x,y>+<y,x>=0$,或 $<x,y>=-<y,x>$.这说明 $<,>$ 是斜对称的.

如果 $\mathrm{char}(F)=2$,那么对于 $\forall \alpha \in F$, $-\alpha=\alpha$,所以对称型和斜对称型的定义是等价的.

设 $\mathrm{char}(F)\neq 2$,如果 $<,>$ 是斜对称的,对于 $\forall x \in V$,我们就有 $<x,x>=-<x,x>$,或 $2<x,x>=0$,则 $<x,x>=0$.因此,$<,>$ 是交错的.

由本定理可知:我们并不需要考虑斜对称型本身,因为斜对称型总等价于对称型或交错型.

例 11.1.2　$V(n,q)$ 上由 $(x_1,\cdots,x_n)\cdot(y_1,\cdots,y_n)=x_1y_1+\cdots+x_ny_n$ 定义的标准内积是对称的,但不是交错的,因为 $(1,0,\cdots,0)\cdot(1,0,\cdots,0)=1\neq 0$.

11.1.2　双线性型的矩阵

定义 11.1.4　如果 $B=\{b_1,\cdots,b_n\}$ 是度量线性空间 V 的一个有序基,那么双线性型 $<,>$ 完全由 $M_B=(a_{i,j})=(<b_i,b_j>)$ 的 $n\times n$ 矩阵决定.

这个矩阵称为关于有序基 B 的双线性型 $<,>$ 的矩阵.

注意到如果 $x=\sum x_ib_i$, $y=\sum y_jb_j$,那么

$$< x,y >= \sum_i \sum_j x_i y_j < b_i, b_j >= \sum_i x_i (\sum_j a_{i,j} y_j) = (x)'_B M_B (y)_B,$$

其中$(x)_B$和$(y)_B$分别是x,y的坐标矩阵.

定义 11.1.5 一个双线性型是**对称**的,当且仅当对于$\forall 1 \leqslant i,j \leqslant n$,这个双线性型的矩阵$M_B = (a_{i,j})$满足$a_{i,j} = a_{j,i}$,即,当且仅当$M_B$是一个对称阵.

同样,一个双线性型是**交错**的,当且仅当这个双线性型的矩阵$M_B = (a_{i,j})$满足$a_{i,i} = 0, a_{i,j} = -a_{j,i} (i \neq j)$.这种矩阵称为是交错的.

现在我们来研究一下关于基变换的双线性型的矩阵是如何表现的.令$C = \{c_1, \cdots, c_n\}$是V的一个有序基.

前面已讲到,第i列是$(c_i)_B$的基矩阵$M_{C,B}$的变换满足$(v)_B = M_{C,B}(v)_C$.因此,

$$< x,y > = (x)'_B M_B (y)_B = [\,(x)'_C M_{C,B}{}'\,] M_B [M_{C,B}(y)_C]$$

$$= (x)_C{}' [M_{C,B}{}' M_B M_{C,B}] (y)_C.$$

所以$M_C = M_{C,B}{}' M_B M_{C,B}$.这就有了以下

定义 11.1.6 两个矩阵$S, T \in M_n(F)$称为是**合同**的,如果存在一可逆矩阵P,使得$S = PTP'$(规定$Q = P'$得$PTP' = Q'TQ$,所以在前面的定义中,我们既可以用PTP',也可以用$Q'TQ$).

定理 11.1.2 如果V上关于有序基$B = \{b_1, \cdots, b_n\}$的双线型的矩阵是$M_B = (< b_i, b_j >)$,那么$< x,y >= (x)'_{C,B} M_B (y)_B$,且如果$C = \{c_1, \cdots, c_n\}$也是$V$的一个有序基,那么$M_C = M'_{C,B} M_B M_{C,B}$.这里$M_{C,B}$是从$C$到$B$,第$i$列为$(c_i)_B$的基矩阵的变换.

因此,如果两个矩阵表示V上的同一个双线性型,那么它们一定是相合的.反之,相合矩阵表示V上的同一个双线性型.假设$B = M_B$表示V上关于有序基B的一个双线性型,且$S = P'TP$,这里P是非奇异的,则存在V的一个有序基X,有性质$P = M_{C,B}$.所以$S = M'_{C,B} M_B M_{C,B}$.因此,关于X的矩阵$S = M_C$与T表示同一个双线性型.

定理 11.1.3 两个矩阵S和T表示V上相同的双线性型$\Leftrightarrow S$和T是合同的.

由于合同矩阵有相同的秩,所以我们可以把一个双线性型的秩定义为任意表示这个双线性型的矩阵的秩.

如果S和T是合同矩阵,那么$\det(S) = \det(PSP') = \det(P)^2 \det(T)$.所以$\det(S)$和$\det(T)$相差一个平方因子.双线性型的**判别式**是在所有有序基的选取之下,表示这个双线性型的矩阵的全体行列式所成的集合.因此,如果对于任意表示双线性型的矩阵S,$\det(S) = d$,那么这个双线性型的判别式是集合$\{r^2 d \mid 0 \neq r \in F\}$.对于一个双线性型的矩阵,通过其判别式通常不会得出太多的结论.

下面,我们将看到对于合同矩阵而言,判别式是一个完全不变式.

11.2　二　次　型

11.2.1　对称双线性型

对称双线性型与另一定义在向量空间的重要函数有着密切联系.

定义 11.2.1　向量空间 V 上的**二次型**是有以下性质的映射 $q: V \to F$:

1) 对于 $\forall r \in F, v \in V, q(rv) = r^2 q(v)$.

2) 映射 $< u,v >_q = q(u+v) - q(u) - q(v)$ 是一个(对称)双线性型.

根据性质 2),每个二次型都定义了一个对称双线性型.

另一方面,如果 $\mathrm{char}(F) \neq 2$, $<,>$ 是 V 上一对称双线性型,那么我们可以由

$$q(x) = \frac{1}{2} < x,x >$$

定义一个二次型 q. 而且,如果 q 是以这样的方式从一个双线性型中定义的,那么与 q 相关的双线性型就是

$$\begin{aligned} < u,v >_q &= q(u+v) - q(u) - q(v) \\ &= \frac{1}{2} < u+v,u+v > - \frac{1}{2} < u,u > - \frac{1}{2} < v,v > \\ &= \frac{1}{2} < u,v > + \frac{1}{2} < v,u > \\ &= < u,v >. \end{aligned}$$

这是一个原始的双线性型,即,映射 $<,> \to q$ 和 $q \to <,>_q$ 是逆映射. 所以在 V 上的对称双线性型和 V 上的二次型之间就存在一一对应的关系.

设 $\mathrm{char}(F) \neq 2$,如果 $B = \{b_1, \cdots, b_n\}$ 是正交几何 V 的一个有序基,且若 V 上对称型的矩阵是 $M_B = (a_{i,j})$,那么对于 $x = \sum x_i b_i$,

$$q(x) = \frac{1}{2} < x,x > = \frac{1}{2}(x)_B^{\mathrm{T}} M_B (x)_B = \sum_{i,j} \frac{1}{2} a_{i,j} x_i x_j.$$

所以 $q(x) = q(x_1, \cdots, x_n)$ 是坐标 x_i 中的一个 2 次齐次多项式.

例 11.2.1　如果 $<,>$ 是 V 上一对称双线性型,则 $q(x) = \frac{1}{2} < x,x >$ 是一个二次型.

例 11.2.2　τ 是一个等距,当且仅当 $q(\tau(v)) = q(v)$,这里 q 是与 V 上的双线性型相关的二次型(其中 $\mathrm{char}(F) \neq 2$).

我们可以定义函数 $\tau: V \to V^*$. 不难看出这个函数是线性的，其核为

$$\{x \in V \mid \varphi_x = 0\} = \{ x \in V \mid 对于\ \forall v \in V, < v,x >= 0\}.$$

因此，如果 V 是非奇异的，那么 τ 的核是零子空间，τ 是单射的. 而且，由于 $\dim(V) = \dim(V^*)$，所以我们可以推出 τ 是满射的. 因此，它是从 V 映到 V^* 上的一个同构映射. 这说明对于 $\forall x \in V$，V 上每个线性泛函都有形式 φ_x. 这样，我们就证明了有限维非奇异度量线性空间上的黎兹表示定理.

定理 11.2.1(黎兹表示定理)　令 V 是一有限维非奇异度量线性空间，且令 $f \in V^*$ 是 V 上一个线性泛函，那么存在唯一的向量 $x \in V$，使得对 $\forall v \in V$，$f(v) = < v,x >$.

11.2.2　正交性

如果 $< x,y >= 0$，那么向量 x 就正交于向量 y，记为 $x \perp y$. 任意正交于本身的非零向量 x 叫做**零向量**，或**迷向向量**.

定理 11.2.2　令 $<,>$ 是 V 上一个双线性型，那么正交性就是一种对称关系，即

$$x \perp y \Leftrightarrow y \perp x,$$

当且仅当 $<,>$ 是对称的或交错的. 因此，在这一情形下，我们就可以有另一种说法：x 和 y 是正交的.

证明　显然，如果 $<,>$ 是对称的，那么 $x \perp y \Leftrightarrow y \perp x$ 成立. 如果 $<,>$ 是交错的，那么它就是斜对称的，所以 $x \perp y \Leftrightarrow y \perp x$ 同样成立.

反过来，假设 $x \perp y \Leftrightarrow y \perp x$ 成立. 对于 V 中的 x，y 和 z，令

$$w = < x,y > z - < x,z > y,$$

那么 $x \perp w$，所以根据假设，$w \perp x$. 而这又等价于 $< x,y >< z,x > - < x,z >< y,x >= 0$.

规定 $y = x$，对于 V 中所有向量 x 和 z，

$$< x,x > (< z,x > - < x,z >) = 0.$$

交换 x 和 z，且乘以 -1，

$$< z,z > (< z,x > - < x,z >) = 0.$$

因此，我们推出，对于 V 中任意向量 u 和 v，如果 $< u,v > \neq < v,u >$，那么 u 和 v 就是零向量. 等价地，如果 u 是非零向量，那么对于 $\forall v \in V$，$< u,v > = < v,u >$.

假设 $<,>$ 不是对称的，那么就存在向量 u 和 v，使得 $< u,v > \neq < v,u >$. 因此

$$< u,u > = < v,v > = 0.$$

我们想要说明对于 $\forall a \in V$，$< a,a >= 0$，从而说明 $<,>$ 是交错的.

因为 u 是零向量,所以 $<u+a,u+a>=<u,a>+<a,u>+<a,a>$.

现在,如果 a 是非零向量,那么对于 $\forall x\in V$, $<a,x>=<x,a>$. 特别地,

$$<a,u>=<u,a>,<a,v>=<v,a>,$$

而且,在 $<x,y><z,x>-<x,z><y,x>=0$ 中规定 $y=a,x=u,z=v$,得

$$<u,a><v,u>-<u,v><a,u>=0,$$

等价于

$$<u,a>(<v,u>-<u,v>)=0.$$

又因为 $<u,v>\neq<v,u>$,所以一定有 $<a,u>=<u,a>=0$. 同样, $<a,v>=<v,a>=0$.

因此,

$$<u+a,u+a>=<u,a>+<a,u>+<a,a>$$

变为 $<u+a,u+a>=<a,a>$. 但

$$<u+a,v>=<u,v>\neq<v,u>=<v,u+a>,$$

所以 $u+a$ 也是零向量,说明 $<a,a>=0$,这与对 a 进行的假设矛盾. 因此, $\forall a\in V$ 都为零,从而 $<,>$ 是交错型.

11.2.3 正交补

如果 W 是度量线性空间 V 的一个子集,那么 W 就有了 V 的度量结构. 这样,我们就称 W 为 V 的一个**子空间**.

定义 11.2.2 如果对于 $\forall s\in W$ 和 $t\in T$, $<s,t>=0$,那么度量线性空间 V 的两个子空间 W 和 T 就是正交的,记为 $W\perp T$. T 称为 W 的**正交补**,记为 $W^{\perp}$,

$$W^{\perp}=\{v\in V\mid v\perp W\}.$$

定义 11.2.3 若 V 是一度量线性空间,那么 $V^{\perp}$ 就叫做 V 的**根**,记为 $\mathrm{Rad}(V)$.

因此, V 是非奇异的,当且仅当 $\mathrm{Rad}(V)=\{0\}$.

如果 W 是 V 的一个子空间,那么 W 的根就是 $\mathrm{Rad}(W)=W\cap W^{\perp}$.

要强调的是,对于任意基域来说,它的正交性性质与我们所熟悉的实基域下的情形有着很大的区别.

例 11.2.3 在实度量线性空间上,我们有 $W\cap W^{\perp}=0$,

而在有限域的度量线性空间上,像下例所说明的那样,甚至会有 $W=W^{\perp}$.

例 11.2.4 不难看出 $V(4,2)$ 的子空间 $S=\{0000,1100,0011,1111\}$ 有性质 $W=W^{\perp}$. $V(4,2)$ 是非奇异的,而其子空间 W 却是奇异的.

定理 11.2.3 如果 W 是非奇异度量线性空间 V 的一个子空间,那么

$$\dim(W) + \dim(W^{\perp}) = \dim(V).$$

证明 对于 $\forall v \in V$,令 φ_v 是 W^* 中由 $\varphi_v: W \to F, \varphi_v(u) = <u,v>$ 定义的线性泛函. 由 $\tau(v) = \varphi_v$ 我们定义映射 $\tau: V \to W^*$. 这一映射是线性的,它的核是

$$\mathrm{Ker}(\tau) = \{v \in V \mid \varphi_v = 0\} = \{ v \in V \mid \text{对于 } \forall u \in W, <u,v> = 0\} = W^{\perp}.$$

而且,根据黎兹定理,限制 $\tau|_W: W \to W^*$ 是满射的,从而 τ 是满射的,即 $\mathrm{Im}(\tau) = W$. 由 $\dim(\mathrm{Im}(\tau)) + \dim(\mathrm{Ker}(\tau)) = \dim(V)$,有 $\dim(W) + \dim(W^{\perp}) = \dim(V)$.

定理 11.2.4 如果 W 是非奇异度量线性空间 V 的一个子空间,那么

1) $(W^{\perp})^{\perp} = W$.

2) $\mathrm{Rad}(W) = W \cap W^{\perp} = \mathrm{Rad}(W^{\perp})$.

下面我们归纳一下度量线性空间中涉及正交性的一些术语(但是,不同作者在使用迷向一词时也会不一样).

定义 11.2.4 令 V 是一度量线性空间.

1) 非零的 $x \in V$ 是**零向量**,或**迷向向量**,如果 $<x,x> = 0$.

2) V 的**根**是 $\mathrm{Rad}(V) = V^{\perp}$.

3) V 是**非奇异的**,或**非退化的**,如果 $V^{\perp} = \{0\}$.

4) V 是**零空间**,如果对于 $\forall u,v \in V$, $<u,v> = 0$,即,若 $V^{\perp} = V$.

5) V 是**迷向空间**,如果 V 中至少含有一个迷向向量.

6) V 是**非迷向空间**,如果 V 中不含有迷向向量.

7) V 是**全迷向空间**,如果 V 中所有向量都是迷向的.

11.2.4 正交直和

定义 11.2.5 令 V 是一度量线性空间. 如果 S 和 T 是 V 的子空间,且有性质 $V = S \oplus T, S \perp T$,那么我们就说 V 是 S 和 T 的**正交直和**,记为 $V = S \odot T$.

在例 11.2.2 中,我们完全可以问一下,对于子空间 S,在什么条件下 $V = S \odot S^{\perp}$ 成立. 以下定理给出了答案.

定理 11.2.5 令 S 是非奇异度量线性空间 V 的子空间,以下等价:

1) S 是非奇异的.

2) $S^{\perp}$ 是非奇异的.

3) $S \cap S^{\perp} = \{0\}$.

4) $V = S + S^{\perp}$.

5) $V = S \odot S^{\perp}$.

证明　根据定理 11.2.4，论断 1)，2) 和 3) 等价. 对于向量空间 V 的任意子空间 S 和 T，有

$$\dim(S+T)=\dim(S)+\dim(T)-\dim(S\cap T),$$

所以

$$\dim(S+S^{\perp})=\dim(S)+\dim(S^{\perp})-\dim(S\cap S^{\perp})=\dim(V)-\dim(S\cap S^{\perp}).$$

这说明论断 3) 与论断 4) 等价，且论断 4) 可推出论断 5). 显然，论断 5) 可推出论断 4)，从而定理获证.

定理 11.2.6　令 V 是一个度量线性空间，那么 $V=\mathrm{Rad}(V)\odot S$，其中 $\mathrm{Rad}(V)$ 为零，S 是非奇异的.

证明　令 S 是 $\mathrm{Rad}(V)$ 的补，即 $V=\mathrm{Rad}(V)\oplus S$. 因为所有向量都正交于 $\mathrm{Rad}(V)$，所以 $\mathrm{Rad}(V)\perp S$，从而 $V=\mathrm{Rad}(V)\odot S$.

现在，如果 $v\in\mathrm{Rad}(S)$，那么 $v\perp S$，从而 $v\perp V$，则 $v\in\mathrm{Rad}(V)\cap S=\{0\}$，即 $v=0$. 因此，$\mathrm{Rad}(S)=\{0\}$，即 S 是非奇异的.

11.3　正交几何的结构

11.3.1　辛几何 — 双曲平面

一般来说，如果 W 是 V 的一个子空间，那么商空间 V/W 不会有 V 的度量结构. 但是，如果 $W=\mathrm{Rad}(V)=V^{\perp}$，那么如下所示，$V/\mathrm{Rad}(V)$ 必有 V 的度量结构. 令

$$<u+\mathrm{Rad}(V),v+\mathrm{Rad}(V)>=<u,v>.$$

为了说明这一内积是唯一确定的，注意到，如果 $u+\mathrm{Rad}(V)=u'+\mathrm{Rad}(V)$，那么 $u=u'+r$，这里 $r\in\mathrm{Rad}(V)$. 因此，

$$\begin{aligned}<u+\mathrm{Rad}(V),v+\mathrm{Rad}(V)>&=<u,v>=<u'+r,v>=<u',v>\\&=<u'+\mathrm{Rad}(V),v+\mathrm{Rad}(V)>.\end{aligned}$$

第二部分的证明类似.

我们考虑一下非奇异辛几何 V. 根据定义，V 中每个向量都为零. 因为 V 假定为是非奇异的，所以取定 $u\in V$，一定存在 $v\in V$，使得 $<u,v>\neq 0$. 考虑基为 $\{u,v\}$ 的二维子空间 H，那么 $<u,u>=<v,v>=0$，$<u,v>=a\neq 0$. 以 $a^{-1}v$ 代替 v，我们可以假设 $<u,v>=1$，$<v,u>=-1$. 这样，度量线性空间的子空间 H(看成是一个度量线性空间) 有关于基 $\{u,v\}$ 的矩阵

$$M_1=\begin{pmatrix}0&1\\-1&0\end{pmatrix}.$$

我们先引入一个定义.

定义 11.3.1 令 V 是一度量线性空间. 如果 $u, v \in V$ 有性质

$$< u,u > = < v,v > = 0,\ < u,v > = 1,$$

那么序偶(u,v) 叫做**双曲偶**,子空间 $H = \text{span}\{u,v\}$ 叫做**双曲平面**. 任意形为

$$H_1 \odot \cdots \odot H_k$$

的空间,其中每一个 H_i 都是一个双曲平面,叫做**双曲空间**.

在正交几何中,如果(u,v) 是一个双曲偶,那么 $< u,v > = 1$,但在辛空间中,

$$< u,v > = -1.$$

现在,我们继续刚才的讨论. 因为 H 是非奇异的,所以 $V = H \odot H^{\perp}$,这里 $H^{\perp}$ 也是非奇异的. 因此,我们可以重复之前在 $H^{\perp}$ 中的构造,得到 V 的形为

$$V = H_1 \odot H_2 \odot \cdots \odot H_k$$

的一个正交分解,这里每个 H_i 都是双曲平面. 这就证明了以下

定理 11.3.1 任意非奇异辛几何 V 都是一个双曲平面,即

$$V = H_1 \odot H_2 \odot \cdots \odot H_k.$$

这里每一 H_i 都是一个双曲平面,因此,存在 V 的一个基,使得双线性型的矩阵为

$$M = \begin{pmatrix} 0 & 1 & & & & & \\ -1 & 0 & & & & & \\ & & 0 & 1 & & & \\ & & -1 & 0 & & & \\ & & & & \ddots & & \\ & & & & & 0 & 1 \\ & & & & & -1 & 0 \end{pmatrix}.$$

特别地,V 的维数是偶数.

推论 11.3.2 任意辛几何都有形式 $V = H_1 \odot H_2 \odot \cdots \odot H_k \odot N$,其中每一 H_i 都是一个双曲空间,N 是一个零空间.

11.3.2 正交几何 — 正交基

与辛几何比起来,正交几何的结构与基域的性质联系更为紧密.

定义 11.3.2 令 V 是一正交几何,V 的基 $B = \{u_1, \cdots, u_n\}$ 称为是**正交的**,如果对于 $i \neq j$, $< u_i, u_j > = 0$.

V 的基 B 是正交的,当且仅当双线性型的矩阵 M_B 是对角阵. 如果 $\text{char}(F) = 2$,那

么任意正交几何都有一个正交基,我们不考虑 V 既是正交几何又是辛几何的情况,因为没有一个非零辛几何能有正交基(根据基的正交性,关于这种基的矩阵的元在对角线外都是 0,而 V 还是一个辛几何,所以对角线上的元也都为 0).

显然,我们不用考虑 V 是零的情形,因为在这一情形下,所有基都是正交的.

我们首先考虑 V 是非奇异的、正交的以及 $\mathrm{char}(F) \neq 2$ 时的情形. 令 $u \in V$ 有性质 $< u,u > \neq 0$. 这种向量一定存在,因为如果不存在的话,那么 V 就是辛的,而且对于 $\mathrm{char}(F) \neq 2$,不存在既是正交的又是辛的非零度量线性空间. 因为子空间 $W = \mathrm{span}\{u\}$ 是非奇异的,所以 $V = W \odot W^{\perp}$,其中 $W^{\perp}$ 是非奇异的和正交的. 因此,可以重复在 $W^{\perp}$ 上的证明,得到 $V = W \odot T \odot T^{\perp}$,其中 W 和 T 是一维子空间. 这样继续下去,得 $V = WS_1 \odot \cdots \odot WS_n$,其中 W_i 由向量 u_i 生成,$< u_i, u_i > \neq 0$. 因此基 $\{u_1, \cdots, u_n\}$ 就是 V 的一个正交基. 这样,由定理 11.2.6 可得任意 $\mathrm{char}(F) \neq 2$ 的正交度量线性空间都有一个正交基.

至于 V 是非奇异的、正交的以及 $\mathrm{char}(F) = 2$ 时的情形. 假设 V 不是辛的,则 V 中存在非零向量 u,从而,和前面一样,$V = S \odot S^{\perp}$.

现在,我们知道如果 $W^{\perp}$ 不是辛的,那么它就是非奇异的和正交的,这样,我们可以再选取一非零向量,并重复上述步骤,直到遇到 V 的一非奇异的、正交的辛空间 T,根据定理 11.2.7,它是双曲平面的一正交和. 因此,

$$V = W_1 \odot \cdots \odot W_k \odot H_1 \odot H_2 \odot \cdots \odot H_m.$$

现在,我们可以证明双线性型的矩阵

$$M = \begin{pmatrix} a & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

与一个对角矩阵合同,其中 $a \neq 0$. 因此,可以用 $T_k \odot T_{k+1} \odot T_{k+2}$ 代替 $W_k \odot H_1$,然后再用这些基向量代替 W_k 和 H_1 的基向量,这里每一直和项都有维数 1.

继续这一步骤,最后得到的 V 是一维子空间所成的一个正交和,所以 V 有一个正交基. 其他满足定理 11.2.6 的情形可以处理一般的(奇异的和非奇异的)情况. 总结如下:

定理 11.3.3　令 V 是一正交几何. 如果当 $\mathrm{char}(F) = 2$ 时,V 不是一个辛几何,那么 V 就有一有序正交基 $B = \{u_1, \cdots, u_k, z_1, \cdots, z_m\}$,使得 $< u_i, u_i > = a_i \neq 0$,$< z_i, z_i > = 0$. 从而 M_B 就有对角形式

$$M_B = \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & a_k & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

这里,对角线上有 k 个 1,m 个零,而且数 k 是双线性型的秩,所以 k 是由 V 唯一确定的.

推论 11.3.4 令 M 是一对称阵.如果 $\mathrm{char}(F)=2$,那么我们设 M 有一个非零元在其对角线上,从而 M 相合于一个对角阵.

11.3.3 正交几何的结构

根据推论 11.3.4,对于 $\mathrm{char}(F)\neq 2$,F 上任意对称阵都相合于一个对角阵.但是,由于两个互异的对角阵是可以相合的,所以我们就不能说对角阵构成相合的一个典范型集.

但是,一个相合的典范型集的确定却依赖于基域的性质.为了更清楚地说明这一点,设 $B=(b_1,\cdots,b_n)$ 是 V 的一个有序正交基,则双线性型的矩阵有以下对角形式:

$$M_B=\begin{bmatrix} a_1 & & & \\ & a_2 & & \\ & & \ddots & \\ & & & a_n \end{bmatrix}.$$

如果 $r_1,\cdots,r_n$ 不为零,那么集合 $C=(r_1b_1,\cdots,r_nb_n)$ 也是 V 的一个有序正交基,且 $< r_ib_i,r_jb_j >= r_ir_j < b_i,b_j >= r_i^2\delta_{i,j}$.

因此关于 C 的双线性型的矩阵为

$$M_C=\begin{bmatrix} r_1^2a_1 & & & \\ & r_2^2a_2 & & \\ & & \ddots & \\ & & & r_n^2a_n \end{bmatrix}, \tag{11.1}$$

从而通过基的一个简单变换,可以给任意对角元乘以 F 中的一个非零平方数.

11.4 有限域上的正交几何

11.4.1 代数闭域

定义 11.4.1 如果 $F[x]$ 中的每一多项式 $p(x)$ 在 F 上都可以分裂为线性因子,称域 F 为**代数闭**的.

例 11.4.1 复数域是代数闭的.但实数域却不是代数闭的.

如果 F 是代数闭的,那么多项式 $x^2-r=0$ 在 F 中有一个解,即 F 的每一元素在 F 中都有一个平方根.

因此，我们可以在上节的 M_C 式中选取 $r_i = \sqrt{a_i}$，这就有了以下

定理 11.4.1　令 V 是代数闭域 F 上一正交几何. 如果当 $\mathrm{char}(F)=2$ 时，V 也不是辛几何，那么 V 就有一有序正交基 $B=(u_1,\cdots,u_k,z_1,\cdots,z_m)$，使得 $<u_i,u_i>=1$，$<z_i,z_i>=0$. 所以 M_B 有以下对角形式：

$$M_B = Z_{k,m} = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

这里，对角线上有 k 个 1，m 个零，而且数 k 是双线性型的秩，所以 k 是由 V 唯一确定的. 特别地，如果 V 也是非奇异的，那么 V 就有一个规范正交基.

定理 11.4.1 的矩阵等价于如下所示的形式：

定理 11.4.2　令 f 是代数闭域 F 上全体 $n\times n$ 对称阵所成的集合. 如果 $\mathrm{char}(F)=2$，我们就把 f 限制为主对角线上至少有一个非零元的全体对称阵所成的集合.

1）对 $\forall k=0,\cdots,n, m=n-k$，$f$ 中任意矩阵都相合于形为 $Z_{k,m}$ 的唯一的矩阵.

2）对于 $k+m=n$，形为 $Z_{k,m}$ 的全体矩阵所成的集合是 f 上合同的典范型集.

3）在 f 上相合的条件之下，矩阵的秩是一个完全不变量.

实数域 $\mathbf{R}$ 不是代数闭的. 但在定理 11.4.2 中，我们可以选取 $r_i=\sqrt{|a_i|}$，使得对角元都为 0，1 或 -1.

定理 11.4.3(西尔维斯特惯性定律)　实数域 $\mathbf{R}$ 上任意正交几何 V 都有一有序正交基

$$B=\{u_1,\cdots,u_k,v_1,\cdots,v_m,z_1,\cdots,z_p\},$$

使得 $<u_i,u_i>=1$，$<v_i,v_i>=-1$，$<z_i,z_i>=0$. 因此，矩阵 M_B 有对角形式

$$M_B = Z_{k,m,p} = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & -1 & & & & & \\ & & & & \ddots & & & & \\ & & & & & -1 & & & \\ & & & & & & 0 & & \\ & & & & & & & \ddots & \\ & & & & & & & & 0 \end{pmatrix}.$$

对角线上有 k 个 1，m 个 -1，p 个 0，而且数 k,m 和 p 是由 V 唯一确定的.

证明 我们只证唯一性. 令

$$P=\operatorname{span}\{u_1,\cdots,u_k\},Q=\operatorname{span}\{v_1,\cdots,v_m\},Z=\operatorname{span}\{z_1,\cdots,z_p\}.$$

那么如果 $v=\sum r_iu_i\in P$，我们就有

$$<v,v>=<\sum r_iu_i,\sum r_ju_j>=\sum_{i,j}r_ir_j<u_i,u_j>=\sum_{i,j}r_ir_j\delta_{i,j}=\sum r_i^2\geqslant 0.$$

所以双线性型 $<,>$ 在 P 上是正定的.

同理，它在 Q 上是负定的，即对于 $\forall v\in Q$，$<v,v>\leqslant 0$.

最后，它在 Z 上为零. 现在假设 C 是一个与 B 的类型相似的有序基，且

$$\overline{P}=\operatorname{span}\{\overline{u_1},\cdots,\overline{u_{\bar{k}}}\},\overline{Q}=\operatorname{span}\{\overline{v_1},\cdots,\overline{v_{\bar{m}}}\},\overline{Z}=\operatorname{span}\{\overline{z_1},\cdots,\overline{z_{\bar{p}}}\},$$

那么 $P\cap\operatorname{span}\{\overline{Q},\overline{Z}\}=\{0\}$. 如果 $v\in P$，那么 $<v,v>\geqslant 0$；如果 $v\in\operatorname{span}\{\overline{Q},\overline{Z}\}$，那么 $<v,v>\leqslant 0$，所以 $v\in P\cap\operatorname{span}\{\overline{Q},\overline{Z}\}$，则 $<v,v>=0$，即 $v=0$. 因此，如果 $\dim(V)=n$，那么

$$\dim(P)+\dim(\operatorname{span}\{\overline{Q},\overline{Z}\})\leqslant\dim(V).$$

即 $p+(n-\overline{p})\leqslant n$，从而 $p\leqslant\overline{p}$. 由对称性，$\overline{p}\leqslant p$，所以 $p=\overline{p}$.

同理，我们推出 $n=\overline{n}$，$z=\overline{z}$.

定理 11.4.4 令 f 是实数域 $\mathbf{R}$ 上全体 $n\times n$ 对称阵所成的集合.

1）对于 $\forall k,m$ 和 $p=n-k-m$，f 中任意矩阵都合同于形 $Z_{k,m,p}$ 的唯一的矩阵.

2）对于 $k+m+p=n$，形为 $Z_{k,m,p}$ 的全体矩阵集是 f 上合同的一个典范型集.

3）在 f 上合同的条件之下，(k,m)，或等价地，$(k+m,k-m)$ 是一个完全不变量. 数 $k+m$ 是双线性型的秩，$k-m$ 叫做双线性型的**符号差**.

11.4.2 有限域

定理 11.4.5 令 F_q 是一个含有 q 个元的有限域.

1）如果 $\operatorname{char}(F_q)=2$，那么 F_q 中每一元都是一个平方数.

2）如果 $\operatorname{char}(F_q)\neq 2$，那么 F_q 中恰有一半非零元是平方数，而且，若 x 是 F_q 中任意非平方数，那么对于 $\forall r\in F_q$，所有非平方数都有形式 r^2x.

证明 1）在任意域 F 中，方程 $x^2=1$ 都有两个解 $x=1$ 和 -1，这两个解是互异的，当且仅当 $\operatorname{char}(F_q)\neq 2$. 令 $F=F_q$，设 F^* 是 F 中全体非零元所成的集合. 考虑 F 中全体非零平方数所成的集合 $(F^*)^2=\{a^2\mid a\in F^*\}$. 注意到对于 $a,b\in F^*$，

$$a^2=b^2\Leftrightarrow(ab^{-1})^2=1\Leftrightarrow ab^{-1}=\pm 1\Leftrightarrow a=\pm b.$$

因此,如果 $\text{char}(F_q)=2$,那么 $a^2=b^2 \Leftrightarrow a=b$. 所以$(F^*)^2=F^*$,论断 1) 获证.

2) 如果 $\text{char}(F_q)\neq 2$,那么映射$(a,-a)\to a^2$ 就是 F 中(互异的) 元素对所成的集合与$(F^*)^2$ 之间的一一对应,所以 $|F^*|=2|(F^*)^2|$.

定义 11.4.2　如果对 $\forall 0\neq r\in F$,都存在向量 v,使得$<v,v>=r$,则称 V 上双线性型是泛的.

定理 11.4.6　令 V 是有限域上一非奇异正交几何,且 $\dim(V)\geqslant 2$,$\text{char}(F)\neq 2$,那么 V 的双线性型就是泛的.

证明　首先设 V 中含有一个零向量 u. 因为 V 是非奇异的,所以 V 中必存在向量 v,使得$\{u,v\}$ 是线性无关的,且 $<u,v>\neq 0$. 令 $w=\alpha u+\beta v$. 对于 $\forall c\neq 0$,我们想要确定 α 和 β,使得 $c=<w,w>=2\alpha<u,v>+\beta^2<v,v>$. 但取定 $\beta=1$ 后,α 就不难解出了. 因此,在这一情形下,$<,>$ 是泛的.

假设 V 中不含有零向量,且$\{u,v\}$ 线性无关,$<u,u>=a\neq 0$,$<v,v>=b\neq 0$,$<u,v>=0$. 令 $w=\alpha u+\beta v$. 我们想要找出 α 和 β,使得 $c=<w,w>=a\alpha^2+b\beta^2$. 用 ac 代替 a,bc 代替 b,同时除以 $c\neq 0$,我们想要说明的是,在任意本征值不为 2 的有限域中,方程 $a\alpha^2+b\beta^2=1$ 总有解(α,β).

如果 a 是一个平方数,那么我们可以规定 $\beta=0$,得到 $\alpha^2=a^{-1}$,或 $\alpha=\sqrt{a^{-1}}$. 同样,如果 b 是一个平方数,那么我们就可以规定 $\alpha=0$,解出 β. 因此,我们假设 a 和 b 都是非平方数.

注意到在 F_q 中,-1 是平方数的和,因为如果 $q=p^n$,那么 F 的本征值就是 p,所以 $-1=1^2+\cdots+1^2$. 这里等式右边有 $p-1$ 个被加数. 因此,任意数 $c\in F$ 都是平方数的和,因为 $4c=(1+c)^2+(-1)(1-c)^2$. 从中,我们推出两个平方数的和不会总是一个平方数,因为如果是这样的话,那么 F_q 中所有元素都将是平方数了,这就与定理 11.4.5 矛盾. 因此 F_q 中存在非零平方数 r 和 s,使得 r^2+s^2 不是一个平方数.

因此,在 F_q 中,a,b 及 r^2+s^2 都不是平方数. 由定理 11.4.5 可知任意两个非平方数的积都是一个平方数,所以推出对于 $\forall u,v\in F$,$b=u^2a$,$r^2+s^2=v^2a$.

规定
$$\alpha=\frac{r}{av},\beta=\frac{s}{uva},$$

得

$$a\alpha^2+b\beta^2=\frac{ar^2}{a^2v^2}+\frac{bs^2}{u^2v^2a^2}=\frac{ar^2u^2+bs^2}{u^2v^2a^2}$$

$$=\frac{b(r^2+s^2)}{u^2v^2a^2}=\frac{b}{u^2a}=1.$$

获证.

11.4.3　有限域上的正交几何

现在我们继续讨论本节的主要内容，首先 $\mathrm{char}(F)=2$ 时的情形.

定理 11.4.7　令 V 是有限域 F 上 $\mathrm{char}(F)=2$ 的一正交几何. 如果 V 不是辛的，那么 V 含有一个有序正交基 $B=\{u_1,\cdots,u_k,z_1,\cdots,z_m\}$，使得 $<u_i,u_i>=1$，$<z_i,z_i>=0$. 所以 M_B 有如下对角形式：

$$M_B=Z_{k,m}=\begin{pmatrix}1&&&&&\\&\ddots&&&&\\&&1&&&\\&&&0&&\\&&&&\ddots&\\&&&&&0\end{pmatrix},$$

其中，对角线上有 k 个 1，m 个零，而且，数 k 是双线性型的秩，由 V 唯一确定.

特别地，如果 V 是非奇异的，那么 V 就有一个规范正交基.

证明　在(11.1)式中，由于 F 中每一元都有一个平方根，所以我们可以取 $r_i=\sqrt{a_i}$.

当 $\mathrm{char}(F)\neq 2$ 时的情形有一点复杂.

定理 11.4.8　令 V 是有限域 F 上 $\mathrm{char}(F)\neq 2$ 的一正交几何，那么存在一个非零数 d 和正交基 $B=\{u_1,\cdots,u_k,z_1,\cdots,z_m\}$，使得 $<u_i,u_i>=1$，对于 $1\leqslant i\leqslant k-1$，$<u_k,u_k>=d$，$<z_i,z_i>=0$. 因此，在这个基中，双线性型的矩阵是

$$M_B=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&1&&&&\\&&&d&&&\\&&&&0&&\\&&&&&\ddots&\\&&&&&&0\end{pmatrix},$$

这个矩阵的秩 k 是由 V 唯一确定的. 如果我们不计 F 中因子的平方数倍，数 d 也是由 V 唯一确定的. 集合 $\{r^2d\mid 0\neq r\in F\}$ 是由 V 唯一确定的，当 V 是非奇异的时，这个集合是双线性型的判别式.

证明　我们知道，存在一个有序正交基 $B=\{u_1,\cdots,u_k,z_1,\cdots,z_m\}$，使得 $<u_i,u_i>=a_i\neq 0$，$<z_i,z_i>=0$，因此，M_B 有对角形式

$$M_B = \begin{pmatrix} a_1 & & & & & \\ & \ddots & & & & \\ & & a_k & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

考虑正交几何 $V_1 = \mathrm{span}\{u_1, u_2\}$，那么 V_1 就是非奇异的，因为 $< u_1, u_2 > \neq 0$，所以限制到 V_1 上的双线性型 $<,>$ 是泛的. 因此，存在 $v_1 \in V_1$，使得 $< v_1, v_1 > = 1$.

由于 $\{u_1, u_2\}$ 是 V_1 的基，所以 $v_1 = ru_1 + su_2$. 如果 $s = 0$，那么我们就构成了有序基

$$B_1 = \{v_1, u_2, \cdots, u_k, z_1, \cdots, z_m\},$$

除了 a_1 的位置是 1 外，关于这个基的双线性型的矩阵与上面 M_{ij} 相同. 如果 $s \neq 0$，那么我们就构成了有序基

$$B_1 = \{v_1, u_1, u_3, \cdots, u_k, z_1, \cdots, z_m\},$$

这里，a_1 的位置为 1，a_2 的位置为 a_1.

在新的有序基 B_1 中，由第二和第三个向量所生成的子空间 V_2 来重复这一步骤，这样继续下去，对于 $1 \leqslant i \leqslant k-1$，可以用 1 代替每一个 a_i.

11.5　维特消去定理

11.5.1　等距

现在我们转而讨论度量线性空间上的等距.

定义 11.5.1　令 V 和 W 是度量线性空间. 我们用同一个记号 $<,>$ 来表示每个空间中的双线性型. 如果对于 V 中所有向量 u, v，有 $< \tau(u), \tau(v) > = < u, v >$，那么双射线性映射 $\tau: V \to W$ 叫做一个**等距**.

如果从 V 到 W 之间存在一个等距，那么就说 V 和 W 是等距的，记为 $V \cong W$.

显然，在合成这一条件下，从 V 到 V 的所有等距所成的集合构成了一个群，叫做 V 的**群**.

如果 V 是一非奇异正交几何，那么从 V 到 V 的等距就叫做**正交变换**.

如果 V 是一非奇异辛几何，那么从 V 到 V 的等距就叫做**辛变换**.

如果 V 是非奇异的，那么等距 $\tau \in L(V)$ 总是单射的（在度量线性空间理论中，都需要等距是双射的，这与特例 —— 实、复内积空间上的情形不一样）.

以下是等距的基本性质：

定理 11.5.1　令 $\tau \in L(V, W)$ 是有限维度量空间 V 和 W 之间的一个线性变换.

1) 令 $B=\{v_1,\cdots,v_n\}$ 是 V 的一个基,那么 τ 是一个等距,当且仅当 τ 是双射的,且对于 $\forall i,j$, $<\tau v_i,\tau v_j>=<v_i,v_j>$.

2) 如果 $\mathrm{char}(F)\neq 2$,那么 τ 是一个等距,当且仅当 τ 是双射的,且对于 $\forall v\in V$, $<\tau(v),\tau(v)>=<v,v>$.

定理 11.5.2 令 $\tau\in L(V)$ 是有限维度量线性空间 V 上一线性算子. 令 $B=\{v_1,\cdots,v_n\}$ 是 V 的一个有序基,且令 M_B 是关于 B 的双线性型的一个矩阵,那么 τ 是一个等距,当且仅当$(\tau)_B^{\mathrm{T}}M_B(\tau)_B=M_B$.

证明 为了方便起见,去掉下标 B, $<x,y>=(x)^{\mathrm{T}}M_B(y)$,且

$$<\tau(x),\tau(y)>=(\tau(x))^{\mathrm{T}}M_B(\tau(y))=(x)^{\mathrm{T}}(\tau)^{\mathrm{T}}M_B(\tau)(y).$$

因此,对于 $\forall x,y\in V$, $<x,y>=<\tau(x),\tau(y)>$,当且仅当对于 $\forall x,y\in V$,

$$(x)^{\mathrm{T}}M_B(y)=(x)^{\mathrm{T}}(\tau)^{\mathrm{T}}M_B(\tau)(y).$$

而上式成立,当且仅当 $M_B=(\tau)^{\mathrm{T}}M_B(\tau)$.

如果 τ 是一个等距,那么$(\tau)_B^{\mathrm{T}}M_B(\tau)_B=M_B$ 成立,我们就可以取行列式,得到

$$\det(M_B)=\det((\tau)_B)^2\det(M_B).$$

因此,如果 V 是非奇异的,那么 $\det(M_B)\neq 0$,从而 $\det((\tau)_B)=\pm 1$.

因为在相似性的条件下,行列式是不变的,所以我们有以下

定义 11.5.2 令 $\tau\in L(V)$ 是一个正交变换. τ 的**行列式**就是任意表示 τ 的矩阵 $(\tau)_B$ 的行列式. 如果 $\det(\tau)=1$,那么 τ 叫做一个**旋转**,如果 $\det(\tau)=-1$,那么 τ 叫做一个**反射**.

由于黎兹表示定理在任意非奇异度量线性空间中都是成立的,所以可以定义线性映射 τ 的伴随 τ^*,即根据条件 $<\tau(v),w>=<v,\tau^*(w)>$.

定理 11.5.3 令 $\tau\in L(V)$ 是有限维非奇异度量线性空间 V 上一线性算子.

1) τ 是一个等距(正交变换),当且仅当 τ 是酉的,即,当且仅当 τ 是双射的,且 $\tau\tau^*=\iota$.

2) 令 τ 是一个等距,如果 $V=S\odot S^{\perp}$,且 S 在 τ 之下是不变的,那么 $S^{\perp}$ 在 τ 之下也是不变的.

证明 2) 由于 S 在 τ 之下是不变的,从而有 $\tau(S)\subset S$. 但是 $\dim(\tau(S))=\dim(S)$,所以 $\tau(S)=S$, $S=\tau^{-1}(S)$. 现在假设 $v\in S^{\perp}$,那么对于 $\forall s\in S$,根据论断 1),以及 $\tau^{-1}(s)\in S$,就有 $<\tau(v),s>=<v,\tau^{-1}(s)>=0$. 因此 $\tau(v)\in S^{\perp}$.

例 11.5.1 如果 $\mathrm{char}(F)\neq 2$,则 $\tau\in L(V,W)$ 是一个等距,当且仅当 τ 是双射的,且对于 $\forall v\in V$, $<\tau(v),\tau(v)>=<v,v>$.

例 11.5.2 令 $B=\{v_1,\cdots,v_n\}$ 是 V 的一个基,则 $\tau\in L(V,W)$ 是一个等距,当且仅当它是双射的,且对于 $\forall i,j$, $<\tau v_i,\tau v_j>=<v_i,v_j>$.

11.5.2　对称

设 V 是 F 上一非奇异度量线性空间，其中 $\mathrm{char}(F)\neq 2$，令 $u\in \mathbf{V}\neq 0$. 考虑线性映射

$$\sigma_u(v)=v-\frac{2<v,u>}{<u,u>}u.$$

定义 11.5.3　如果 σ_u 有以下性质：

1) σ_u 是一个等距；2) $\sigma_u(u)=-u$；3) 对于 $\forall x\in(\mathrm{span}\{u\})^{\perp}$，$\sigma_u(x)=x$，那么我们就称 σ_u 为由 u 确定的**对称**.

因为 $V=\mathrm{span}\{u\}\odot(\mathrm{span}\{u\})^{\perp}$，所以线性映射 σ_u 是由性质 2) 和 3) 唯一确定的. 而且，$\sigma_u=-\iota\,|_{<u>}\odot\iota\,|_{<u>^{\perp}}$，其中 $<u>$ 是由向量 u 所生成的子空间.

我们还需要对称的以下性质：

定理 11.5.4　令 V 是域 F 上一非奇异正交几何，其中 $\mathrm{char}(F)\neq 2$，u 和 v 是 V 的非零向量，且 $<u,u>=<v,v>\neq 0$，那么就存在对称 σ，使得 $\sigma(u)=v$ 或 $\sigma(u)=-v$.

证明　假设 $u+v$ 不等于 0，那么对称 σ_{u+v} 是有意义的，且 $\sigma_{u+v}(u+v)=-(u+v)$. 而且，由于 $<u-v,u+v>=0$，所以 $\sigma_{u+v}(u-v)=u-v$. 把这两个等式相结合，就得到我们想要的 $\sigma_{u+v}(u)=-v$.

如果 $u+v$ 等于 0，那么 $u-v$ 一定不为 0，否则 u 就会为 0. 因此，对称 σ_{u+v} 是有意义的，而且，$\sigma_{u-v}(u-v)=-(u-v)$，$\sigma_{u-v}(u+v)=u+v$. 从这两个等式中就得出了我们想要的 $\sigma_{u-v}(u)=v$.

定理 11.5.5　令 V 是域 F 上一非奇异正交几何，其中 $\mathrm{char}(F)\neq 2$，那么任意正交变换 $\tau: V\to V$ 是 V 上对称的积.

证明　我们在 $d=\dim(V)$ 上运用归纳法，假设对于 $d-1$ 定理成立，令 $\dim(V)=d$，选取一非零向量 $u\in V$，由于 $<\tau(u),\tau(u)>=<u,u>$，所以我们可以运用定理 11.5.3 推出 V 上对称 σ 的存在性，使得 $\sigma(\tau(u))=\varepsilon u$，其中 $\varepsilon=\pm 1$. 因为 σ 是一个对称，所以如果 $x\in<u>^{\perp}$，那么 $<\sigma\tau(x),u>=<\sigma\tau(x),\varepsilon\sigma\tau(u)>=\varepsilon<x,u>=0$. 从而 $\sigma\tau(<u>^{\perp})\subset<u>^{\perp}$. 因此，$<u>$ 和 $<u>^{\perp}$ 在 $\sigma\tau$ 之下都是不变的. 根据应用在 $d-1$ 维空间 $<u>^{\perp}$ 上的归纳假设，

$$\sigma\tau\,|_{<u>^{\perp}}=\sigma_{w_1}\cdots\sigma_{w_k}. \tag{11.2}$$

其中 $w_i\in<u>^{\perp}$，σ_{w_i} 是 $<u>^{\perp}$ 上由 $\sigma_{w_i}(v)=v-\dfrac{2<v,w_i>}{<w_i,w_i>}w_i$ 所定义的对称. 这也是 V 上的一个对称，其中 $\sigma_{w_i}(u)=u$. 因此，把 σ_{w_i} 看成定义在 V 上之后，就有

$$\sigma_{w_1}\cdots\sigma_{w_k}(u)=u=\varepsilon\sigma\tau(u). \tag{11.3}$$

我们要区分两种情况. 如果 $\varepsilon=1$，那么(11.2) 式和(11.3) 式就表明在 $<u>$ 和 $<u>^{\perp}$ 上，$\sigma\tau=\sigma_{w_1}\cdots\sigma_{w_k}$，则在 V 上，$\sigma\tau=\sigma_{w_1}\cdots\sigma_{w_k}$.

最后,由于对称是它本身的逆,所以 $\tau = \sigma\sigma_{w_1} \cdots \sigma_{w_k}$.

另一方面,如果 $\varepsilon = -1$,那么由(11.3)式得,

$$\sigma_u\sigma_{w_1}\cdots\sigma_{w_k}(u) = \sigma_u(u) = -u = \sigma\tau(u),$$

由于 σ_u 固定 $\langle u \rangle^{\perp}$ 中的所有向量,$\langle u \rangle^{\perp}$ 在 $\sigma\tau$ 之下是不变的,对于 $x \in \langle u \rangle^{\perp}$,由(11.2),有

$$\sigma_u\sigma_{w_1}\cdots\sigma_{w_k}(x) = \sigma_u\sigma\tau\,|_{\langle u \rangle^{\perp}}(x) = \sigma\tau\,|_{\langle u \rangle^{\perp}}(x).$$

因此,在这一情形下,$\sigma\tau = \sigma_u\sigma_{w_1} \cdots \sigma_{w_k}$,从而 $\tau = \sigma\sigma_u\sigma_{w_1} \cdots \sigma_{w_k}$.

11.5.3　维特消去定理

现在,我们要讨论正交几何的主要结论之一 —— 维特消去定理.

定理 11.5.6 (维特消去定理)　令 V 是域 F 上一非奇异正交几何,其中 $\mathrm{char}(F) \neq 2$. 设 $V = S \odot S^{\perp} = T \odot T^{\perp}$,其中 S 和 T 是非奇异的,那么 $S \cong T, S^{\perp} \cong T^{\perp}$.

证明　令 $\tau: S \to T$ 是一个等距. 我们在 $\dim(S)$ 上运用归纳法. 首先假设 $\dim(S) = 1, S = \mathrm{span}\{s\}$,那么 $T = \mathrm{span}\{\tau(s)\}$,$\langle \tau(s), \tau(s) \rangle = \langle s, s \rangle$.

根据定理 11.5.3,存在对称 σ,使得 $\sigma(s) = \varepsilon\tau(s)$,这里 $\varepsilon = \pm 1$. 因此,σ 是 V 的使得 $\sigma(S) = T$ 的一个等距. 由此得出

$$x \in S^{\perp} \Leftrightarrow \langle x, s \rangle = 0 \Leftrightarrow \langle \sigma(x), \sigma(s) \rangle = 0 \Leftrightarrow \langle \sigma(x), \tau(s) \rangle = 0 \Leftrightarrow \sigma(x) \in T^{\perp}.$$

所以限制 $\sigma\,|_{S^{\perp}}$ 是从 $S^{\perp}$ 到 $T^{\perp}$ 的一个等距,这说明 $S^{\perp} \cong T^{\perp}$.

现在设对于 $\dim(S) < k$,定理成立,令 $\dim(S) = k$,令 $\tau: S \to T$ 是一个等距. 由于 S 是非奇异的,所以可以选取一非零向量 $s \in S$,记为 $S = \mathrm{span}\{s\} \odot U$,这里 U 是非奇异的,而且 $T = \mathrm{span}\{\tau(s)\} \odot \tau(U)$. 因此,

$$V = \mathrm{span}\{s\} \odot U \odot S^{\perp};\ V = \mathrm{span}\{\tau(s)\} \odot \tau(U) \odot T^{\perp}.$$

对于一维的情形,我们可以运用维特定理推出 $U \odot S^{\perp} \cong \tau(U) \odot T^{\perp}$.

假设 $\sigma: U \odot S^{\perp} \to \tau(U) \odot T^{\perp}$ 是一个等距. 由于 $\sigma(U \odot S^{\perp}) = \sigma(U) \odot \sigma(S^{\perp})$,所以我们有 $\sigma(U) \odot \sigma(S^{\perp}) = \tau(U) \odot T^{\perp}$. 但是 $\sigma(U) \cong \tau(U)$,又因为

$$\dim(\sigma(U)) = \dim(U) < \dim(S),$$

所以由归纳假设即得 $S^{\perp} \cong \sigma(S^{\perp}) \cong T^{\perp}$.

11.6　维特扩张定理

11.6.1　维特扩张定理

设 V 和 V' 是非奇异正交几何,$\sigma: V \to V'$ 是一个等距. 同样,设 U 是 V 的一非奇异

子空间，$\tau: U \to \tau(U) \subset V'$ 是一个等距. 我们想要说明 τ 能扩充成 V 上的一个等距. 因为 U 是非奇异的，所以 $\tau(U)$ 也是非奇异的，从而我们推出

$$V' = \tau(U) \odot \tau(U)^{\perp} = \sigma(U) \odot \tau(U^{\perp}).$$

因为 $\tau(U) \cong U \cong \sigma(U)$，所以维特消去定理就蕴含着 $\tau(U)^{\perp} \cong \sigma(U^{\perp})$. 如果 $\upsilon: \sigma(U^{\perp}) \to \tau(U)^{\perp}$ 是一个等距，那么积 $\upsilon\sigma: U^{\perp} \to \tau(U)^{\perp}$ 也是一个等距，从而 $\bar{\tau} = \tau \odot \upsilon\sigma: U \odot U^{\perp} \to V'$ 是一个有性质 $\bar{\tau}|_U = \tau$ 的等距. 这就是 τ 的扩张.

我们要说明 U 是非奇异的这一假设是不需要的. 首先要说明的是 V 的任意子空间 U 都可以嵌入到一个非奇异子空间中，并且 U 上任意等距都可以扩充到这一非奇异子空间中，那么我们就可以在之前所描述的非奇异的情形中寻求帮助.

定理 11.6.1　令 V 是 F 上一个非奇异正交几何，其中 $\mathrm{char}(F) \neq 2$. 令 U 是 V 的一个子空间，记作 $U = \mathrm{Rad}(U) \odot W$，其中 W 是非奇异的. 设 $B = \{b_1, \cdots, b_k\}$ 是 $\mathrm{Rad}(U)$ 的一个基，那么存在向量 $\{z_1, \cdots, z_k\}$，使得

1) (b_i, z_i) 是一双曲对，从而空间 $H_i = \mathrm{span}\{b_i, z_i\}$ 是一双曲平面；

2) U 包含在非奇异空间 $H_1 \odot \cdots \odot H_k \odot W$ 中.

而且，如果 $\tau: U \to \tau(U) \subset V'$ 是一个等距，这里 V' 是非奇异的，那么存在等距 $\bar{\tau}: V \to V'$，使得 $\bar{\tau}|_U = \tau$.

证明　我们通过在 $k = \dim(\mathrm{Rad}(U))$ 上进行归纳来证论断 1) 和论断 2). 当 $k = 0$ 时，显然成立. 令 $k = 1$，因而 $B = \{b_1\}$ 是 $\mathrm{Rad}(U)$ 的一个基. 我们想要找到 $z_1 \in V$，使得 (b_1, z_1) 是一个双曲对，即

$$< b_1, b_1 > = < z_1, z_1 > = 0, (b_1, z_1) = 1. \tag{11.4}$$

这里，令 $H_1 = \mathrm{span}\{b_1, z_1\}$ 后，我们就有 $H_1 \cap W = \{0\}, H_1 \perp W$.

如果我们能找到 $z_1 \in W^{\perp}$，使得上式成立，则若 $x = rb_1 + sz_1 \in H_1 \cap W$，那么

$$0 = < rb_1 + sz_1, z_1 > = r.$$

所以 $x = sz_1 \in W \cap W^{\perp} = \{0\}$. 由于 W 是非奇异的，所以 $H_1 \cap W = \{0\}$.

由于 $b_1, z_1 \in W^{\perp}$ 可得 $H_1 \perp W$，因此我们只要找到 $z_1 \in W^{\perp}$，使得 $< b_1, b_1 > = < z_1, z_1 > = 0, (b_1, z_1) = 1$ 成立. 因为 $b_1 \in W^{\perp}, \mathrm{Rad}(W^{\perp}) = \mathrm{Rad}(W) = \{0\}$，所以一定存在向量 $x \in W^{\perp}$，使得 $< b_1, x > \neq 0$. 我们规定 $z_1 = rb_1 + sx$，说明存在 r 和 s，使 (11.4) 式成立，即使得 $1 = < b_1, z_1 > = < b_1, rb_1 + sx > = s < b_1, x >$，且

$$0 = < z_1, z_1 > = < rb_1 + sx, rb_1 + sx > = 2rs < b_1, x > + s^2 < x, x >.$$

因为 $< b_1, x > \neq 0$，所以从第一个等式中可以解出 s，从第二个中解出 r. 因此，存在向量 z_1，使得 (b_1, z_1) 是一个双曲对，从而当 $k = 1$ 时，论断 1) 和论断 2) 成立.

根据归纳法，假设对于 $\dim(\mathrm{Rad}(U)) < k$，论断 1) 和论断 2) 成立.

令 $\dim(\mathrm{Rad}(U))=k$，那么 $U=\mathrm{Rad}(U)\odot W=\mathrm{span}\{b_k\}\odot\mathrm{span}\{b_1,\cdots,b_{k-1}\}\odot W$.

令 $U_0=\mathrm{span}\{b_1,\cdots,b_{k-1}\}\odot W$，则 $\mathrm{Rad}(U_0)=\mathrm{span}\{b_1,\cdots,b_{k-1}\}$. 因为 $b_k\in U_0^{\perp}$，且

$$\mathrm{Rad}(U_0^{\perp})=\mathrm{Rad}(U_0)=\mathrm{span}\{b_1,\cdots,b_{k-1}\}$$

不含有 b_k，所以和上面一样，我们可以推出存在向量 $x\in U_0^{\perp}$，使得 $(b_k,x)\neq 0$，且存在向量 $z_k\in U_0^{\perp}$，使得 (b_k,z_k) 是一个双曲对. 令 $H_k=\mathrm{span}\{b_k,z_k\}$. 由于 $H_k\subset U_0^{\perp}$，所以 $U_0\subset H_k^{\perp}$，又因为 $H_k^{\perp}$ 是非奇异的，$\dim(\mathrm{Rad}(U_0))=k-1$，所以可以在 $H_k^{\perp}$ 的子空间 U_0 上应用归纳假设. 因此 $H^{\perp}$ 中存在双曲平面 $H_1,\cdots,H_{k-1}$，使得

$$U_0\subset\overline{U_0}=H_1\odot\cdots\odot H_{k-1}\odot W.$$

又因为 $\overline{U_0}\subset H_k^{\perp}$，所以

$$H_k\perp\overline{U_0},U\subset H_1\odot\cdots\odot H_k\odot W.$$

最后一部分，设 $\tau\colon U\to\tau(U)\subset V'$ 是一个等距，这里 V' 是非奇异的，则

$$U=\langle b_1\rangle\odot\cdots\odot\langle b_k\rangle\odot W\subset H_1\odot\cdots\odot H_k\odot W.$$

这里 $\langle b_i\rangle$ 是由 b_i 所生成的子空间，$H_i=\mathrm{span}\{b_i,z_i\}$. 由于 τ 是一个等距，所以

$$\tau(U)=\langle\tau(b_1)\rangle\odot\cdots\odot\langle\tau(b_k)\rangle\odot\tau(W).$$

现在，令 $\bar{\tau}(b_i)=\tau(b_i)$，且对于 $\forall w\in W,\bar{\tau}(w)=\tau(w)$. 只要选取 $\bar{\tau}(z_i)$，使得 $\bar{\tau}$ 是一个等距.

最后，将 1) 应用到含有子空间 $\tau(U)$ 的非奇异空间 V' 上. 由于 $\{\tau(b_1),\cdots,\tau(b_k)\}$ 是 $\mathrm{Rad}\{\tau(U)\}$ 的一个基，所以我们推出存在向量 $w_i\in V'$，使得

$$\tau(U)\subset K_1\odot\cdots\odot K_k\odot W',$$

这里 $K_i=\mathrm{span}\{\tau(b_i),w_i\}$ 是 V' 的双曲平面. 因此，如果我们令 $\bar{\tau}(z_i)=w_i$，就不难得出 $\bar{\tau}$ 是一个等距了.

定理 11.6.2(维特扩张定理)　令 V,V' 是域 F 上的等距非奇异正交几何，其中 $\mathrm{char}(F)\neq 2$. 设 U 是 V 的一非奇异子空间，$\tau\colon U\to U'\subset V'$ 是一个等距，那么 τ 可以扩充到 V 上，即存在等距 $\bar{\tau}\colon V\to V'$，使得 $\bar{\tau}|_U=\tau$.

下面我们考虑维特扩张定理的应用. 令 V 是域 F 上一非奇异正交几何，其中 $\mathrm{char}(F)\neq 2$. 设 U,U' 是 V 的极大零子空间. 我们想要说明 $\dim(U)=\dim(U')$.

如果 $\dim(U)\leqslant\dim(U')$，那么存在同构 $\tau\colon U\to\tau(U)\subset U'$，因为 U 和 U' 是零子空间，所以这是一个等距. 因此由扩张定理就有扩充 τ 的等距 $\bar{\tau}\colon V\to V$ 的存在.

特别地，$\bar{\tau}^{-1}(U')$ 是含有 U 的零空间，则 $\bar{\tau}^{-1}(U')=U$，这说明 $\dim(U)=\dim(U')$.

定理 11.6.3　令 V 是域 F 上一非奇异正交几何，$\mathrm{char}(F)\neq 2$，则 V 的所有极大零子空间都有相同维数，叫做 V 的**维特指数**，记为 $w(V)$.

11.6.2　极大双曲子空间

由于双曲空间完全是由它的维数决定的(如果我们不计等距的话)，所以我们就想知道非奇异正交几何的极大双曲子空间(在辛几何的情形中，如果 V 是非奇异的，那么 V 就是双曲的). 我们用 W 来表示一个双曲空间，维数为 $2k$ 的双曲空间表示为 W_{2k}，因此

$$W_{2k} = H_1 \odot \cdots \odot H_k.$$

这里每个 H_i 都是一个双曲平面.

二维空间是一个双曲平面，当且仅当它是非奇异的，且包含一个零向量.

设 V 是迷向的，即 V 中含有一个零向量. 如果 U_k 是 V 的维数为 k 的一个非空零子空间，那么 $\mathrm{Rad}(U_k) = U_k$，从而我们可以推出

$$U_k \subset W_{2k} = H_1 \odot \cdots \odot H_k.$$

这里 H_i 由双曲对 (x_i, y_i) 所生成. 因此，任意零子空间 U_k 都包含在双曲空间 W_{2k} 中，且 $\dim(W_{2k}) = 2\dim(U_k)$. 这说明 V 的维特指数最多为 $\dim(V)/2$.

另一方面，设 $W_{2k} = H_1 \odot \cdots \odot H_k$ 是 V 的一双曲空间，且 H_i 由双曲对 (x_i, y_i) 所生成，那么集合 $B = \{x_1, \cdots, x_k\}$ 是无关的，因为如果 $r_1 x_1 + \cdots + r_k x_k = 0$，那么从中可以得出，对于 $\forall j$，

$$0 = \langle r_1 x_1 + \cdots + r_k x_k,\ y_j \rangle = r_j \langle x_j, y_j \rangle = r_j.$$

而且，由于对于 $\forall i, j$，$\langle x_i,\ y_j \rangle = 0$，所以子空间 $U_k = \mathrm{span}\{B\}$ 是一个 k 维零空间，则 V 中任意双曲空间 W_{2k} 都含有一个零空间 U_k，则若 W_{2m} 是 V 的一个极大双曲子空间，那么 $m \leqslant w(V)$. 而且，由于 V 必含有零空间 $U_{w(V)}$，所以 V 也一定含有一个维数为 $2w(V)$ 的双曲子空间. 换句话说，V 的双曲子空间的最大维数为 $2w(V)$.

现在，设 $W = H_1 \odot \cdots \odot H_k$，且 $E = K_1 \odot \cdots \odot K_m$ 是 V 的极大双曲子空间，同时设 H_i 由双曲对 (u_i, v_i) 生成，K_i 由双曲对 (x_i, y_i) 生成.

我们想要说明 $\dim(W) = \dim(E)$. 设 $\dim(W) \leqslant \dim(E)$，考虑由条件 $\tau(u_i) = x_i$，$\tau(v_i) = y_i$ 定义的向量空间单一同态 τ：$W \to \tau(W) \subset E$. 根据定理 11.5.1，τ 是一个等距，所以 $W \cong \tau(W)$. 因此，维特扩张定理说明扩充 τ 的等距 $\bar{\tau}$：$V \to V$ 存在.

特别地，$\bar{\tau}^{-1}(E)$ 是一个含有 W 的双曲空间，所以 $\bar{\tau}^{-1}(E) = W$，这说明 $\dim(E) = \dim(W)$. 这样我们就证明了 V 的所有极大双曲子空间都有相同的维数 $2w(V)$. 假设 W 是 V 的一个极大双曲子空间，由于双曲空间是非奇异的，所以 $V = W \odot W^{\perp}$，那么 $W^{\perp}$ 就是非迷向的，即 $W^{\perp}$ 不含有零向量. 为了说明这一点，反过来设 $x \in \mathbf{W}^{\perp}$ 是一个零向量. 因为存在零子空间 $U \subset W$，使得 $\dim(U) = \dim(W)/2 = w(V)$，所以零空间 $U' = \mathrm{span}\{U, E\}$ 有维数 $w(V) + 1$，这与维特指数的意义矛盾. 因此，$W^{\perp}$ 中不含有零向量.

反之，设 W_{2k} 是双曲的，$V=W_{2k}\odot W_{2k}^{\perp}$，$W_{2k}^{\perp}Z$ 是非迷向的，那么 W_{2k} 是极大双曲的，因为如果 W_{2k} 不是极大双曲，那么 W_{2k} 就真包含于双曲子空间 W_{2m} 中，我们记为 $W_{2m}=W_{2k}\odot E$. 这里 E 是 W_{2m} 中 W 的正交补. E 必是一个维数为 $2(m-k)$ 的双曲空间. 因为对于任意双曲空间 $W_{2(m-k)}$，必有 $W_{2m}=W_{2k}\odot W_{2(m-k)}$，所以根据维特消去定理，$E\cong W_{2(m-k)}$. 由于 $W_{2(m-k)}$ 含有一个零向量，所以由上式可知存在一个零向量 $x\in W_{2k}^{\perp}$，这与假设矛盾. 因此，W_{2k} 是极大的.

定理 11.6.4　令 V 是 F 上一非奇异正交几何，$\mathrm{char}(F)\neq 2$，那么 V 的极大双曲子空间都有维数 $2w(V)$，这里 $w(V)$ 是 V 的维特指数，且 $V=W\odot S$.

这里 W 是 V 的一个极大双曲子空间，或者 $W=\{0\}$，如果 V 中不含有零向量，且 S 是 V 的一个非迷向子空间，那么 S 中不含有零向量.

域 F 上任意正交几何 V，其中 $\mathrm{char}(F)\neq 2$，都可以记为 $\mathrm{Rad}(V)\odot W\odot S$. 这里 $\mathrm{Rad}(V)$ 是一个零空间，W 是一个双曲空间，S 是非迷向的.

第 12 章　希尔伯特空间

前面我们研究了实、复内积空间的基本性质. 当时,我们不太关注所讨论的空间是有限维还是无限维的. 但是,正如我们在第 9 章中所讨论的,向量空间上内积、距离的存在,产生了许多关于收敛的新问题. 在本章中,我们讨论距离空间、希尔伯特空间的概念,从而研究实、复内积空间的收敛性,更详细地讨论最佳逼近的问题. 我们的目标是从希尔伯特空间 H 的一闭子空间 S 中,找出任意向量 x 的最佳逼近的一种清晰的表达式.

12.1　距离空间上的收敛性

12.1.1　距离空间

距离空间并不是一种代数结构,而是用来模拟距离的抽象性质的.

定义 12.1.1　**距离空间**即(M,d),这里 M 是一非空集,$d: M\times M\to \mathbf{R}$ 是一个实值函数,叫做 M 上的**距离**,$d(x,y)$ 读作"从 x 到 y 的距离". 它有以下性质.

1)(正定性)　对于 $\forall\, x,y\in M, d(x,y)\geqslant 0$.

2)(对称性)　对于 $\forall\, x,y\in M, d(x,y)=d(y,x)$.

3)(三角不等式)　对于 $\forall\, x,y,z\in M, d(x,y)\leqslant d(x,z)+d(z,y)$.

例 12.1.1　在由 $d(x,y)=\begin{cases}0, & 若\ x=y;\\ 1, & 若\ x\neq y\end{cases}$ 定义的离散距离之下,任意非空集 M 是一个距离空间.

例 12.1.2　1) 对于 $x=(x_1,\cdots,x_n), y=(y_1,\cdots,y_n)$,在由

$$d(x,y)=\sqrt{(x_1-y_1)^2+\cdots+(x_n-y_n)^2}$$

定义的距离之下,集合 $\mathbf{R}^n$ 是一个距离空间,叫做 $\mathbf{R}^n$ 上的**欧几里得距离**.

在距离 $d_1(x,y)=|x_1-y_1|+\cdots+|x_n-y_n|$ 之下,$\mathbf{R}^n$ 也是一个距离空间. 当然$(\mathbf{R}^n,d)$ 和$(\mathbf{R}^n,d_1)$ 是两个不同的距离空间.

2) 在酉距离 $d(x,y)=\sqrt{|x_1-y_1|^2+\cdots+|x_n-y_n|^2}$ 下,集合 $\mathbf{C}^n$ 是一个距离

空间,这里 $x=(x_1,\cdots,x_n)$,$y=(y_1,\cdots,y_n)$ 在 $\mathbf{C}^n$ 中.

例 12.1.3　1) 在距离 $d(f,g)=\sup\limits_{x\in[a,b]}|f(x)-g(x)|$ 之下,$[a,b]$ 上全体实值(或复值)连续函数所成的集合 $C[a,b]$ 是一个距离空间.我们称这种度量为**上确界距离**.

2) 在距离 $d_1(f(x),g(x))=\int_a^b|f(x)-g(x)|\mathrm{d}x$ 之下,$[a,b]$ 上全体实值(或复值)连续函数所成的集合 $C[a,b]$ 是一个距离空间.

例 12.1.4　许多重要的序列空间都是距离空间.与 $\mathbf{x}=(x_n)$,$\mathbf{y}=(y_n)$ 一样,我们将会用粗体罗马字母来表示序列.

1) 在由 $d(\mathbf{x},\mathbf{y})=\sup\limits_n|x_n-y_n|$ 定义的距离之下,全体有界实数序列所成的集合 L_R^∞ 是一个距离空间.含有同一个距离的全体有界复数序列所成的集合 L_C^∞ 也是一个距离空间.我们通常用 L^∞ 表示这些空间.

2) 对于 $p\geqslant 1$,令 L^p 是使得 $\sum\limits_{n=1}^{\infty}|x_n|^p<\infty$ 的实(复)数的全体序列 $\mathbf{x}=(x_n)$ 所成的集合.由 $\|\mathbf{x}\|_p=(\sum\limits_{n=1}^{\infty}|x_n|^p)^{1/p}$,定义 $\mathbf{x}$ 的 **p 范数**.那么在距离 $d(\mathbf{x},\mathbf{y})=\|\mathbf{x}-\mathbf{y}\|_p=(\sum\limits_{n=1}^{\infty}|x_n-y_n|^p)^{1/p}$ 之下,L^p 是一个距离空间.

定理 12.1.1(赫尔德不等式)　令 $p,q\geqslant 1$,$p+q=pq$.如果 $\mathbf{x}\in L^p$,$\mathbf{y}\in L^q$,那么 $\mathbf{xy}=(x_n,y_n)\in L^1$,且 $\|\mathbf{xy}\|_1\leqslant\|\mathbf{x}\|_p\|\mathbf{y}\|_q$,即

$$\sum_{n=1}^{\infty}|x_ny_n|\leqslant(\sum_{n=1}^{\infty}|x_n|^p)^{1/p}(\sum_{n=1}^{\infty}|y_n|^q)^{1/q}.$$

它的特例($p=q=2$)是柯西 — 施瓦茨不等式

$$\sum_{n=1}^{\infty}|x_ny_n|\leqslant\sqrt{\sum_{n=1}^{\infty}|x_n|^2}\sqrt{\sum_{n=1}^{\infty}|y_n|^2}.$$

定理 12.1.2(闵可夫斯基不等式)　对于 $p>1$,如果 $\mathbf{x},\mathbf{y}\in L^p$,那么 $\mathbf{x}+\mathbf{y}=(x_n+y_n)\in L^p$,且 $\|\mathbf{x}+\mathbf{y}\|_p\leqslant\|\mathbf{x}\|_p+\|\mathbf{y}\|_p$,即

$$(\sum_{n=1}^{\infty}|x_n+y_n|^p)^{1/p}\leqslant(\sum_{n=1}^{\infty}|x_n|^p)^{1/p}+(\sum_{n=1}^{\infty}|y_n|^p)^{1/p}.$$

定义 12.1.2　如果 M 是 d 之下的一个距离空间,那么在 d 到 $S\times S$ 的限制之下,M 的任意非空子集 S 也是一个距离.距离空间 S 叫做 M 的一个**子空间**.

12.1.2　拓扑空间

定义 12.1.3　令 M 是一距离空间,$x_0\in M$,r 是一个正实数.

1）中心在 x_0，半径为 r 的**开球**是 $B(x_0,r)=\{x\in M \mid d(x,x_0)<r\}$.

2）中心在 x_0，半径为 r 的**闭球**是 $\overline{B}(x_0,r)=\{x\in M \mid d(x,x_0)\leqslant r\}$.

3）中心在 x_0，半径为 r 的**球面**是 $S(x_0,r)=\{x\in M \mid d(x,x_0)=r\}$.

定义 12.1.4　距离空间 M 的子集 S 称为是**开**的，如果 S 的每一点都是一个完全包含在 S 中的开球的中心. 更具体地说，S 是开的，如果对于 $\forall x\in S$，都存在一个 $r>0$，使得 $B(x,r)\subset S$.

开球就是开集，闭球就是闭集.

如果 $x\in M$，那么我们把任意包含 x 的开集 S 称为 x 的**开领域**.

同样，一个集合是开的，当且仅当它包含由其每个点所成的开领域.

下面一个例子说明一个集合既是开的又是闭的，或既不是开的，也不是闭的，这两种情形都可能出现.

例 12.1.5　在距离空间 $\mathbf{R}$ 中，开球正好是开区间 $B(x_0,r)=(x_0-r,x_0+r)$.

闭球是闭区间 $\overline{B}(x_0,r)=[x_0-r,x_0+r]$.

对于 $a<b$，考虑半开区间 $S=(a,b]$. 这个集合不是开的，因为它不含有中心在 $b\in S$ 的开球，也不是闭的，因为它的补 $S'=(-\infty,a]\cup(b,\infty)$ 不是开的（它不含有关于 a 的开球）. 同样，和整个空间 $\mathbf{R}$ 一样，空集既是开的，又是闭的（可以说明在 $\mathbf{R}$ 中，空集和 $\mathbf{R}$ 是唯一的两个既是开集又是闭集的集合）.

开集和闭集属于数学的一个分支 —— 拓扑. 我们的目的并不是要详细地讨论这两种集合，但是为了更清楚地说明相关概念，我们有以下结论：

定理 12.1.3　距离空间 M 中全体开集所成的集合 Ω 有以下性质：

1）$\varnothing\in\Omega, M\in\Omega$.

2）如果 $S,T\in\Omega$，那么 $S\cap T\in\Omega$.

3）如果 $\{S_i \mid i\in K\}$ 是开集所成的任意集合，那么 $\bigcup\limits_{i\in K} S_i\in\Omega$.

定义 12.1.5　X 是一非空集. X 的子集所成的集合 Ω 叫做 X 的**拓扑**，如果满足

1）$\varnothing\in\Omega, X\in\Omega$.

2）如果 $S,T\in\Omega$，那么 $S\cap T\in\Omega$.

3）如果 $\{S_i \mid i\in K\}$ 是 Ω 中集合所成的任意集合，那么 $\bigcup\limits_{i\in K} S_i\in\Omega$.

我们称 Ω 的子集为**开集**，(X,Ω) 为**拓扑空间**.

距离空间 M 中的开集（与我们之前对其下的定义一样）构成 M 的一个拓扑，叫做由距离所**诱导**的拓扑.

拓扑空间是最普通的空间，在其中我们可以定义诸如收敛、连续性之类的概念，这就是这些概念称为拓扑概念的原因所在. 但是，由于我们将要涉及的拓扑是由距离诱导的，所以对于那些我们直接就要用的关于拓扑性质的定义，一般会根据距离来进行定义.

12.1.3　距离空间上的收敛性

定义 12.1.6　如果$\lim\limits_{n\to\infty}d(x_n,x)=0$,称距离空间$M$中序列$\{x_n\}$ **收敛**于$x\in M$,记作$\{x_n\}\to x$. 等价地$\{x_n\}\to x$,如果对于$\forall\varepsilon>0$,存在$N>0$,使得$n>N$,则有$d(x_n,x)<\varepsilon$. 或等价地,$n>N$,则有$x_n\in B(x,\varepsilon)$. 在这一情形下,x叫做序列$\{x_n\}$的**极限**. M是一距离空间,S是M的子集,那么S中的序列,即为所有项都在S中的序列.

定理 12.1.4　令M是一距离空间. 子集$S\subset M$是闭的,当且仅当如果(x_n)是S中的一个序列,且$(x_n)\to x$,那么$x\in S$. 或子集S是闭的,如果S在取序列极限时是闭的.

证明　设S是闭集,令$(x_n)\to x$,对于$\forall n,x_n\in S$. 设$x\notin S$,那么由于$x\in S^c$,S^c是开集,所以存在$\varepsilon>0$,使得$x\in B(x,\varepsilon)\subset S^c$,则$B(x,\varepsilon)\cap\{x_n\}=\varnothing$. 这与$(x_n)\to x$矛盾. 因此$x\in S$.

反之,设S在取极限时是闭的. 我们来说明S^c是开的. 令$x\in S^c$,设没有一个关于x的开球是含在S^c中的. 对于$n=1,2,\cdots$考虑开球$B(x,1/n)$. 因为对于$\forall n$,没有一个这样的开球是包含在S^c中的,所以存在$x_n\in S\cap B(x,1/n)$. 显然$(x_n)\to x$,从而$x\in S$. 但x不能既在S中又在S^c中,所以关于x的任意球在S^c中,则S^c是开的,从而S是闭的.

定义 12.1.7　令S是距离空间M的任意子集. S的**闭包**,记为$cl(S)$,是S的最小闭集.

定义 12.1.8　令S是距离空间M的一非空子集. 元素$x\in M$称为S的一个**极限点**,或**聚点**,如果每一个中心在x的开球除了x本身外,还与S交于一点. 我们用$l(S)$表示S的全体极限点所成的集合.

定理 12.1.5　令S是距离空间M的一非空子集.

1) 元素$x\in M$是S的一个极限点$\Leftrightarrow S$中存在序列$\{x_n\}$,对于$\forall n,x_n\neq x$,且$\{x_n\}\to x$.

2) S是闭的$\Leftrightarrow l(S)\subset S$,即,$S$是闭的$\Leftrightarrow S$包含其所有的极限点.

3) $cl(S)=S\cup l(S)$.

4) 元素x在$cl(S)$中$\Leftrightarrow M$中存在序列(x_n),使得$(x_n)\to x$.

证明　1) 首先假设$x\in cl(S)$. 对于$\forall n$,存在点$x_n\neq x$,使得$x_n\in S(x,1/n)\cap S$. 因此,$d(x_n,x)<1/n$. 所以$\{x_n\}\to x$.

反之,设$(x_n)\to x$,其中$x\neq x_n\in S$. 如果$B(x,r)$是中心在x的任意球,那么存在$\forall N$,使得当$n>N$时有$x_n\in B(x,r)$. 因此,对于任意中心在x的球$B(x,r)$,都存在一个点$x_n\neq x$,使得$x_n\in S\cap B(x,r)$. 因此,x是S的一个极限点.

2) 如果S是闭的,那么根据论断1),$\forall x\in l(S)$都是S中序列(x_n)的极限,所以也就一定在S中. 因此,$l(S)\subset S$. 反之,如果$l(S)\subset S$,那么S是闭的. 因为如果(x_n)是S中任意序列,$(x_n)\to x$,那么就存在两种可能性:

ⅰ. 对于 $\forall n, x_n = x$，此时，$x = x_n \in S$；

ⅱ. 对于 $\forall n, x_n \neq x$，此时，$(x_n) \to x$，则 $x \in l(S) \subset S$.

上述无论哪种情况都有 $x \in S$，所以在取极限时 S 是闭的，也就是说 S 是闭的.

3) 显然 $S \subset T = S \cup l(S)$. 为了说明 T 是闭的，我们来说明 T 包含其全体极限点. 令 $x \in l(T)$，从而存在序列 $(x_n) \in T$，使得 $x_n \neq x$，$(x_n) \to x$. 当然，每个 x_n 或者在 S 中，或者是 S 的一个极限点. 我们必须说明 $x \in T$，即，x 或者在 S 中，或者是 S 的一个极限点.

反之，设 $x \notin S, x \notin l(S)$，那么存在球 $B(x,r)$，使得 $B(x,r) \cap S \neq \varnothing$. 由于 $(x_n) \to x$，所以一定存在 $x_n \in B(x,r)$. 因为 x_n 不可能在 S 中，所以它一定是 S 的一个极限点. 如果 $d(x_n, x) = d < r$，那么考虑球 $B\left(x_n, \dfrac{r-d}{2}\right)$. 这个球完全包含在 $B(x, r)$ 中，而且一定包含 S 的一个元素 y，因为它的中心 x_n 是 S 的一个极限点. 但这样 $y \in S \cap B(x, r)$，矛盾. 从而，$x \in S$ 或 $x \in l(S)$. 无论是哪种情况，$x \in T = S \cup l(S)$，所以 T 是闭的. 因此，T 是闭的且包含 S，所以 $cl(S) \subset T$.

另一方面，$T = S \cap l(S) \subset cl(S)$，所以 $cl(S) = T$.

4) 如果 $x \in cl(S)$，那么就存在两种可能性. 如果 $x \in S$，那么对于 $\forall x$，常数列 $\{x_n\}$，其中 $x_n = x$，是 S 中收敛于 x 的序列. 如果 $x \notin S$，那么 $x \in l(S)$. 所以 S 中存在序列 $\{x_n\}$，使得 $x_n \neq x$，$(x_n) \to x$. 无论哪种情况，S 中都存在一个收敛于 x 的序列.

反之，如果 S 中存在序列 (x_n)，使得 $\{x_n\} \to x$，那么或者对于 $\forall n, x_n = x$，此时 $x \in S \subset cl(S)$，或者对于 $\forall n, x_n \neq x$，此时 $x \in l(S) \subset cl(S)$.

12.2　距离空间的稠密与连续

12.2.1　稠密性

下面的概念说明子集 $S \subset M$“任意接近”于 M 中的每一点.

定义 12.2.1　距离空间 M 的子集 S 在 M 中是**稠密的**，如果 $cl(S) = M$. 一个距离空间称为是**可分的**，如果它含有一个可数稠密子集.

因此，M 的子集 S 是稠密的，如果每一关于任意点 $x \in M$ 的开球都至少含 S 的一个点.

当然，任意距离空间都含有一个稠密子集，即这个空间本身.

但是，如下例，并不是每个距离空间都含有一个可数稠密子集.

例 12.2.1　1) 实直线 $\mathbf{R}$ 是可分的，因为有理数 $\mathbf{Q}$ 构成了一个可数稠密子集. 同样，$\mathbf{R}^n$ 是可分的，因为集合 $\mathbf{Q}^n$ 是可数和稠密的.

2) 复平面 $\mathbf{C}$ 是可分的,同样,对于 $\forall n$,$\mathbf{C}^n$ 也是可分的.

3) 离散距离空间是可分的,当且仅当它是可数的.

例 12.2.2 空间 L^∞ 不是可分的,L^∞ 是全体有界实数(或复数)序列所成的集合,其距离为

$$d(x,y)=\sup_n |x_n-y_n|.$$

为了说明这一空间是不可分的,考虑全体二进制序列所成的集合 S

$$S=\{(x_n) \mid \text{对于 } \forall i, x_i=0 \text{ 或 } 1\}.$$

这一集合在 $\mathbf{N}$ 的全体子集所成的集合的一一对应中,所以是不可数的(它有基数 $2^{\aleph_0}>\aleph_0$). 现在,S 中每个序列都是有界的,从而也就在 L^∞ 中. 而且,如果 $x\neq y\in L^\infty$,那么这两个序列肯定至少有一个位置是不同的,因此 $d(x,y)=1$.

换句话说,我们有 L^∞ 的一个不可数子集 S,使得任意两个互异元之间的距离为 1,则球 $\left\{B\left(s,\frac{1}{3}\right) \mid s\in S\right\}$ 所成的不可数集是互不相交的.

因此,没有一个集合可以与每个球都相交,这就说明在 L^∞ 中,没有一个可数集可以是稠密的.

例 12.2.3 对于 $p\geqslant 1$,距离空间 L^p 是可分的. 对于 $\forall n>0$,形为 $s=(q_1,\cdots,q_n,0,\cdots)$ 的全体序列所成的集合 S 是一个可数集,这里 q_i 是有理数.

我们来说明它在 L^p 中是稠密的.

$\forall x\in L^p$ 满足 $\sum_{n=1}^{\infty}|x_n|^p<\infty$. 对于 $\forall \varepsilon>0$,存在一个 N,使 $\sum_{n=N+1}^{\infty}|x_n|^p<\frac{\varepsilon}{2}$.

因为有理数在 $\mathbf{R}$ 中稠密,所以存在有理数 q_i,使得对于 $\forall i=1,\cdots,N$, $|x_i-q_i|^p<\frac{\varepsilon}{2N}$. 如果 $s=(q_1,\cdots,q_N,0,\cdots)$,那么

$$d(x,s)^p=\sum_{n=1}^{N}|x_n-q_n|^p+\sum_{n=N+1}^{\infty}|x_n|^p<\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.$$

这说明 S 中存在一元素,它任意接近于 L^p 中任意元素. 因此,S 在 L^p 中是稠密的,从而 L^p 是可分的.

12.2.2 连续性

在研究无限维内积空间上的线性算子时,连续性起着重要作用.

定义 12.2.2 令 $f: M\to M'$ 是从距离空间 (M,d) 到距离空间 (M',d') 上的一个函数. 如果对于 $\forall \varepsilon>0$,存在 $\delta>0$,使得 $d(x,x_0)<\delta \Rightarrow d'(f(x),f(x_0))<\varepsilon$,或等价地,$f(B(x_0,\delta))\subset B(f(x_0),\varepsilon)$,我们就说 f 在 $x_0\in M$ 处是**连续**的. 一个函数是连续的,如

果它在每一 $x_0 \in M$ 处都是连续的.

我们可运用收敛的概念描述两个距离空间之间的函数的连续性.

定理 12.2.1　函数 $f: M \to M'$ 是连续的,当且仅当 $\{x_n\}$ 是 M 中收敛于 $x_0 \in M$ 的序列时,序列 $\{f(x_n)\}$ 就收敛于 $f(x_0)$,即由 $\{x_n\} \to x_0$ 可得 $\{f(x_n)\} \to f(x_0)$.

证明　设 f 在 x_0 处是连续的,令 $\{x_n\} \to x_0$,那么给定 $\varepsilon > 0$,由 f 的连续性可得,存在 $\delta > 0$,使得 $f(B(x_0,\delta)) \subset B(f(x_0),\varepsilon)$. 因为 $\{x_n\} \to x_0$,所以存在 $N > 0$,使得对于 $n > N$, $x_0 \in B(x_0,\delta)$,从而由 $n > N$ 可得 $f(x_n) \in B(f(x_0),\varepsilon)$. 因此,$\{f(x_n)\} \to f(x_0)$.

反之,设 $\{x_n\} \to x_0$,则 $\{f(x_n)\} \to f(x_0)$. 假设 f 在 x_0 处是不连续的,那么存在 $\varepsilon > 0$,使得对于 $\forall \delta > 0, f(B(x_0,\delta)) \underset{\neq}{\subset} B(f(x_0),\varepsilon)$. 因此,对于 $\forall n > 0$,

$$f(B(x_0,1/n)) \underset{\neq}{\subset} B(f(x_0),\varepsilon).$$

所以,通过选取有性质 $x_n \in B(x_0,1/n)$,但 $f(x_n) \notin B(f(x_0),\varepsilon)$ 的每一项 x_n,我们可以构造序列 $\{x_n\}$,因此 $\{x_n\} \to x_0$;但是 $f(x_n)$ 不收敛于 $f(x_0)$. 这一矛盾说明在 x_0 处,f 一定是连续的.

定理 12.2.2　令 (M,d) 是一距离空间. 如果 $\{x_n\} \to x$, $(y_n) \to y$,则 $d(x_n,y_n) \to d(x,y)$.

证明　有 $d(x_n,y_n) - d(x,y) | \leqslant d(x_n,x) + d(y_n,y)$.

当 $n \to \infty$ 时上式右边趋向于 0,所以 $d(x_n,y_n) \to d(x,y)$.

12.2.3　完全性

定义 12.2.3　距离空间 M 上序列 $\{x_n\}$ 是一个**柯西序列**,如果对于 $\forall \varepsilon > 0$,都存在 $N > 0$,使得由 $n,m > N$ 可得 $d(x_n,x_m) < \varepsilon$.

任意收敛序列都是柯西序列. 当逆命题成立时,这个空间就称为是完全的.

定义 12.2.4　令 M 是一距离空间.

1) M 称为是**完全的**,如果 M 中每个柯西序列都在 M 中收敛.

2) M 的子空间 S 是**完全的**,如果作为一个距离空间,S 是完全的.

因此,S 是完全的,如果 S 中每一柯西序列 (s_n) 都收敛于 S 中的一个元素.

定理 12.2.3　令 M 是一距离空间.

1) M 的任意完全子空间都是闭的.

2) 如果 M 是完全的,那么 M 的子空间 S 是完全的 $\Leftrightarrow S$ 是闭的.

证明　1) 设 S 是 M 的一个完全子空间. 令 $\{x_n\}$ 是 S 中的一个序列,这里 $\{x_n\} \to x \in M$,那么 (x_n) 是 S 的一个柯西序列,由于 S 是完全的,所以 $\{x_n\}$ 一定收敛于 S 中的一个元素. 因为序列的极限是唯一的,所以 $x \in S$. 因此,S 是闭的.

2) 首先设 S 是完全的,那么由论断 1),S 是闭的.

反之,设 S 是闭的,令 $\{x_n\}$ 是 S 的一个柯西序列. 由于 $\{x_n\}$ 在完全空间 M 中也是一个柯西序列,所以它一定收敛于任意 $x \in M$. 但由于 S 是闭的,所以 $\{x_n\} \to x \in M$,从而 S 是完全的.

例 12.2.4 距离空间 $\mathbf{R}$ 是一个完全空间. 同样,复数 $\mathbf{C}$ 是完全的.

例 12.2.5 欧几里得空间 $\mathbf{R}^n$ 和酉空间 $\mathbf{C}^n$ 都是完全的. 我们来证 $\mathbf{R}^n$ 是完全空间. 设 $\{x_k\}$ 是 $\mathbf{R}^n$ 中的一个柯西序列,这里 $x_k = (x_{k,1}, \cdots, x_{k,n})$,因此,$k, m \to \infty$ 时,

$$d(x_k, x_m)^2 = \sum_{i=1}^{n} (x_{k,i} - x_{m,i})^2 \to 0.$$

所以,对于每个坐标位置 i,$(x_{k,i} - x_{m,i})^2 \leqslant d(x_k, x_m)^2 \to 0$. 这说明在 $\mathbf{R}$ 中,i 个坐标所成的序列 $(x_{k,i})$,$k = 1, 2, \cdots$ 是一个柯西序列. 由于 $\mathbf{R}$ 是完全的,所以必有 $k \to \infty$ 时,$(x_{k,i}) \to y_i$. 如果 $y = (y_1, \cdots, y_n)$,那么 $k \to \infty$ 时,$d(x_k, y)^2 = \sum\limits_{i=1}^{n} (x_{k,i} - y_i)^2 \to 0$. 所以 $\{x_n\} \to y \in \mathbf{R}^n$. 从而,$\mathbf{R}^n$ 是完全的.

例 12.2.6 $[a,b]$ 上,距离为 $d(f,g) = \sup\limits_{x \in [a,b]} |f(x) - g(x)|$ 的全体实值(或复值)连续函数所成的距离空间 $(C[a,b], d)$ 是完全的. 注意到关于 d 的极限是 $[a,b]$ 上的一致极限,即 $d(f_n, f) \to 0$,当且仅当对于 $\forall \varepsilon > 0$,都存在 $N > 0$,使得当 $n > N$ 时,对于 $\forall x \in [a,b]$,$|f_n(x) - f(x)| \leqslant \varepsilon$.

现在,令 $\{x_n\}$ 是 $(C[a,b], d)$ 的一柯西序列. 因此,对于 $\forall \varepsilon > 0$,都存在 $N > 0$,使得,当 $m, n > N$ 时,对于 $\forall x \in [a,b]$,$|f_n(x) - f_m(x)| \leqslant \varepsilon$. 这说明,对于 $\forall x \in [a,b]$,序列 $(f_n(x))$ 都是实(或复)数所成的一个柯西序列,所以 $(f_n(x))$ 收敛. 因此,由 $f(x) = \lim\limits_{n \to \infty} f_n(x)$,我们可以在 $[a,b]$ 上定义函数 f.

在 $|f_n(x) - f_m(x)| \leqslant \varepsilon$ 中,令 $m \to \infty$,当 $n > N$ 时,对于 $\forall x \in [a,b]$,$|f_n(x) - f(x)| \leqslant \varepsilon$. 因此,$f_n(x)$ 一致收敛于 $f(x)$. 由于连续函数的一致极限是连续的,所以 $f(x) \in C[a,b]$. 从而,$(f_n(x)) \to f(x) \in C[a,b]$,所以 $(C[a,b], d)$ 是完全的.

例 12.2.7 距离空间 L^∞ 是完全的. 为了说明这一点,设在 L^∞ 中,$\{x_n\}$ 是一柯西序列,这里 $x_n = (x_{n,1}, x_{n,2}, \cdots)$,那么,对于每个坐标位置 i,当 $n, m \to \infty$ 时,

$$|x_{n,i} - x_{m,i}| \leqslant \sup_j |x_{n,j} - x_{m,j}| \to 0. \tag{12.1}$$

因此,对于 $\forall i$,i 个坐标所成的序列 $(x_{n,i})_{n=1,2,\cdots}$ 是 $\mathbf{R}$(或 $\mathbf{C}$) 中的一个柯西序列. 由于 $\mathbf{R}$(或 $\mathbf{C}$) 是完全的,所以对于每一坐标位置 $i = 1, 2, \cdots, n \to \infty$ 时,$(x_{n,i}) \to y_i$. 我们想要说明 $y = (y_i) \in L^\infty$,$(x_n) \to y$. 在(12.1)式中,令 $m \to \infty$,当 $n \to \infty$ 时,$\sup\limits_j |x_{n,j} - y_j| \to 0$,所以对于 $\forall n, j$,$|x_{n,j} - y_j| < 1$,从而对于 $\forall j$,$|y_j| < 1 + |x_{n,j}|$. 但由于 $x_n \in l^\infty$,所以 (x_n) 是一个有界序列,从而 (y_j) 也是一个有界序列,即 $y = (y_j) \in L^\infty$. 由

$\sup_j |x_{n,j}-y_j| \to 0$ 得 $\{x_n\} \to y$,所以 L^∞ 是完全的.

例 12.2.8　距离空间 L^p 是完全的.

令 $\{x_n\}$ 是 L^p 中一个柯西序列,$x_n=(x_{n,1},x_{n,2},\cdots)$,那么对于每个坐标位置 i,

$$|x_{n,i}-x_{m,i}|^p \leqslant \sum_{j=1}^{\infty} |x_{n,j}-x_{m,j}|^p = d(x_n,x_m)^p \to 0.$$

这说明,i 个坐标所成的序列 $\{x_{n,i}\}_{n=1,2,\cdots}$ 是 **R**(或 **C**) 中的一个柯西序列. 由于 **R**(或 **C**) 是完全的,所以当 $n\to\infty$ 时,$\{x_{n,i}\}\to y_i$.

对于 $\forall \varepsilon>0$,都存在 N,使得对于 $\forall r>0$,由 $n,m>N$ 可得

$$\sum_{i=1}^{r} |x_{n,i}-x_{m,i}|^p \leqslant \varepsilon.$$

令 $m\to\infty$,对于 $\forall r>0$,由 $n>N$ 可得

$$\sum_{i=1}^{r} |x_{n,i}-y_i|^p \leqslant \varepsilon.$$

令 $r\to\infty$,对于 $\forall n>N$,有 $\sum_{i=1}^{\infty} |x_{n,i}-y_i|^p<\varepsilon$,则 $\{x_n\}-y\in L^p$,所以 $y=y-\{x_n\}+\{x_n\}\in L^p$,而且 $\{x_n\}\to y$.

完全性的性质在内积空间理论中起着重要作用. 使得诱导距离空间为完全空间的内积空间叫做**希尔伯特空间**.

12.3　距离空间的完全化

12.3.1　距离空间中柯西序列的等价类

两个距离空间之间保持距离的函数叫做等距.

定义 12.3.1　令 (M,d) 和 (M',d') 是距离空间. 函数 $f\colon M\to M'$ 叫做**等距**,如果对于 $\forall x,y\in M, d'(f(x),f(y))=d(x,y)$.

如果 $f\colon M\to M'$ 是从 M 到 M' 的一个双射等距,我们就说 M 和 M' 是**等距的**,记作 $M\approx M'$.

定理 12.3.1　令 $f\colon(M,d)\to(M',d')$ 是一个等距,那么

1) f 是单射的.

2) f 是连续的.

3) $f^{-1}\colon f(M)\to M$ 也是一个等距,从而也是连续的.

证明　1) 我们注意到 $f(x)=f(y) \Leftrightarrow d'(f(x),f(y))=0 \Leftrightarrow d(x,y)=0 \Leftrightarrow x=y$.

2) 令在 M 中,$(x_n)\to x$,那么当 $n\to\infty$ 时,$d'(f(x_n),f(x))=d(x_n,x)\to 0$,所以

$f(x_n) \to f(x)$，从而 f 是连续的.

3) 因为 $d(f^{-1}(f(x)), f^{-1}(f(y))) = d(x,y) = d'(f(x), f(y))$，所以 f^{-1}：$f(M) \to M$ 是一个等距.

定理 12.3.2 令 (M,d) 是任意距离空间，那么存在完全距离空间 (M',d')，以及等距 τ：$M \to \tau(M) \subset M'$，使得 $\tau(M)$ 在 M' 中是稠密的. 距离空间 (M',d') 叫做 (M,d) 的一个**完全化**. 而且，在不计双射等距的前提下，(M',d') 是唯一的.

证明 首先我们把 M 中的序列 $\{x_n\}$ 看作函数 f：$N \to M$，这里 $f(n) = x_n$（N 是自然数集合）. 证明的基本思想是使 M' 中的元素成为 M 中柯西序列的等价类. 因此，令 $CS(M)$ 表示 M 中全体柯西序列所成的集合.

如果 $f, g \in CS(M)$，那么当 $n \to \infty$ 时，项 $f(n)$ 就会趋向于一个数，从而项 $g(n)$ 也趋向于一个数. 因此，当 $n \to \infty$ 时，$d(f(n), g(n))$ 应该趋向于一个有限极限.

事实上，当 $n, m \to \infty$ 时，

$$| d(f(n),g(n)), d(f(m),g(m)) | \leqslant d(f(n),f(m)) + d(g(n),g(m)) \to 0.$$

所以 $d(f(n),g(n))$ 是实数所成的一个柯西序列，则 $\lim\limits_{n\to\infty} d(f(n),g(n)) < \infty$，即极限存在，并且是有限的. 由 $d'(f,g) = \lim\limits_{n\to\infty} d(f(n),g(n))$，我们想要在集合 $CS(M)$ 上定义距离 d'. 但是，对于互异的序列 f 和 g，可能存在 $\lim\limits_{n\to\infty} d(f(n),g(n)) = 0$，所以不能在集合 $CS(M)$ 上定义距离.

因此，我们就要根据 $f \sim g \Leftrightarrow \lim\limits_{n\to\infty} d(f(n),g(n)) = 0$，在 $CS(M)$ 上定义一个等价关系. 令 $\overline{CS(M)}$ 是柯西序列中全体等价类所成的集合，对 $\overline{f}, \overline{g} \in \overline{CS(M)}$，定义

$$d'(\overline{f},\overline{g}) = \lim_{n\to\infty} d(f(n),g(n)),$$

这里 $f \in \overline{f}, g \in \overline{g}$. 设 $f' \in \overline{f}, g' \in \overline{g}$，那么由于 $f' \sim f, g' \sim g$，所以当 $n \to \infty$ 时，

$$| d(f'(n),g'(n)) - d(f(n),g(n)) | \leqslant d(f'(n),f(n)) + d(g'(n),g(n)) | \to 0.$$

因此，

$$f' \sim f, g' \sim g \Rightarrow \lim_{n\to\infty} d(f'(n),g'(n)) = \lim_{n\to\infty} d(f(n),g(n)) \Rightarrow d'(f',g') = d'(f,g),$$

这说明 d' 是唯一确定的.

下面说明 d' 是一个距离，如果 f, g, h 是柯西序列，那么 $d(f(n),g(n)) \leqslant d(f(n),h(n)) + d(h(n),g(n))$. 取极限后，得

$$\lim_{n\to\infty} d(f(n),g(n)) \leqslant \lim_{n\to\infty} d(f(n),h(n)) + \lim_{n\to\infty} d(h(n),g(n)).$$

所以，$d'(\overline{f},\overline{g}) \leqslant d'(\overline{f},\overline{h}) + d'(\overline{h},\overline{g})$. 对于 $\forall x \in M$，考虑常数柯西序列 $\{x\}$，这里对于 $\forall n, \{x\}(n) = x$. 由 $\tau(x) = \overline{\{x\}}$ 定义的映射 τ：$M \to M'$ 是一个等距，因为

$$d'(\tau(x),\tau(y)) = d'(\overline{\{x\}},\overline{\{y\}}) = \lim_{n\to\infty} d(\{x\}(n),\{y\}(n)) = d(x,y).$$

而且 $\tau(M)$ 在 M' 中是稠密的. 我们可以用一个常数列来逼近 M 中的任意柯西序列.

特别地,令 $\overline{f} \in M'$. 由于 $f \in \overline{f}$ 是一个柯西序列,所以对于 $\forall \varepsilon > 0$,存在一个 N,使得 $n,m \geqslant N \Rightarrow d(f(n),f(m)) < \varepsilon$. 现在,对于常数列 $\{f(N)\}$,

$$d'(\overline{\{f(N)\}},\overline{f}) = \lim_{n\to\infty} d(f(N),f(n)) \leqslant \varepsilon,$$

所以 $\tau(M)$ 在 M' 中是稠密的.

12.3.2　(M',d') 的完全性

设 $\overline{f}_1,\overline{f}_2,\cdots$ 是 M' 中一柯西序列. 在 M 中找出一柯西序列 g,使得当 $k\to\infty$ 时,$d'(\overline{f}_k,\overline{g}) = \lim_{n\to\infty} d(f_k(n),g(n)) \to 0$. 由于 $\overline{f}_k \in M'$,$\tau(M)$ 在 M' 中是稠密的,所以存在一常数列 $\{c_k\}$,使得 $d'(\overline{f}_k,\overline{\{c_k\}}) = \lim_{n\to\infty} d(f_k(n),c_k) < 1/k$. 令 g 是由 $g(k)=c_k$ 所定义的一个序列. 这是 M 中一个柯西序列,因为当 $k,j\to\infty$ 时,

$$\begin{aligned} d(c_k,c_j) = d'(\overline{\{c_k\}},\overline{\{c_j\}}) &\leqslant d'(\overline{\{c_k\}},\overline{f}_k) + d'(\overline{f}_k,\overline{f}_j) + d'(\overline{f}_j,\overline{\{c_j\}}) \\ &\leqslant 1/k + d'(\overline{f}_k,\overline{f}_j) + 1/j \to 0. \end{aligned}$$

为了说明 $\overline{f}_k$ 收敛于 $\overline{g}$,注意到

$$d'(\overline{f}_k,\overline{g}) \leqslant d'(\overline{f}_k,\overline{\{c_k\}}) + d'(\overline{\{c_k\}},\overline{g}) < 1/k + \lim_{n\to\infty} d(c_k,g(n)) = 1/k + \lim_{n\to\infty} d(c_k,c_n).$$

由于 g 是一个柯西序列,则 $\forall \varepsilon > 0$,存在 N,当 $k,n \geqslant N$ 时,有 $d(c_k,c_n) < \varepsilon$.

特别地,当 $k \geqslant N$ 时,有 $\lim_{n\to\infty} d(c_k,c_n) \leqslant \varepsilon$. 所以当 $k \geqslant N$ 时,有 $d'(\overline{f}_k,\overline{g}) \leqslant 1/k + \varepsilon$,这就得到了我们想要的 $\overline{f}_k \to g$.

12.3.3　唯一性

最后,如果 (M',d') 和 (M'',d'') 都是 (M,d) 的完全化,那么 $M' \approx M''$. 我们有双射等距 τ: $M \to \tau(M) \subset M'$ 和 σ: $M \to \sigma(M) \subset M''$. 因此,映射 $\rho = \sigma\tau^{-1}$: $\tau(M) \to \sigma(M)$ 是一个从 $\tau(M)$ 映到 $\sigma(M)$ 的双射等距,这里 $\tau(M)$ 在 M' 中是稠密的(图 12-1).

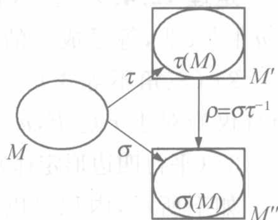


图 12-1

我们的目的是要说明 ρ 可以扩充为一个从 M' 到 M'' 的双射等距 $\overline{\rho}$.

令 $x \in M$,那么 $\tau(M)$ 中存在序列 $\{a_n\}$,使得 $\{a_n\} \to x$.

由于$\{a_n\}$是$\tau(M)$中一柯西序列，所以$\{\rho\{a_n\}\}$是$\sigma(M)\subset M''$中一柯西序列，又因为M''是完全的，所以对于$\forall y\in M''$，$\{\rho\{a_n\}\}\to y$. 下面定义$\bar{\rho}(x)=y$.

为了说明$\bar{\rho}$是唯一确定的，设$\{a_n\}\to x$，$\{b_n\}\to x$，其中这两个序列都在$\tau(M)$中，那么当$n\to\infty$时，$d''(\rho\{a_n\},\rho\{b_n\})=d'(a_n,b_n)\to 0$. 所以$\{\rho\{a_n\}\}$和$\{\rho\{b_n\}\}$收敛于$M''$中的同一个元，则$\bar{\rho}(x)$与$\tau(M)$中收敛于$x$的序列的选取无关，因此，$\bar{\rho}$是唯一确定的. 而且，如果$a\in\tau(M)$，那么常数列$\{a\}$收敛于$a$，所以$\bar{\rho}(a)=\lim\rho\{a\}=\rho(a)$，这说明$\bar{\rho}$是$\rho$的一个扩张.

设$\{a_n\}\to x$，$\{b_n\}\to y$，那么$\{\rho\{a_n\}\}\to\bar{\rho}(x)$，$\{\rho\{b_n\}\}\to\bar{\rho}(y)$，由于$d''$是连续的，所以$d''(\bar{\rho}(x),\bar{\rho}(y))=\lim\limits_{n\to\infty}d''(\rho\{a_n\},\rho\{b_n\})=\lim\limits_{n\to\infty}d'(a_n,b_n)=d'(x,y)$则$\bar{\rho}$是一个等距，因为$\sigma(M)=\mathrm{Im}(\rho)\subset\mathrm{Im}(\bar{\rho})$，所以如果$\mathrm{Im}(\bar{\rho})$是闭的，那么由$\sigma(M)$在$M''$中是稠密的，可推出$\mathrm{Im}(\bar{\rho})=M''$.

所以，设$\{\bar{\rho}\{x_n\}\}$是$\mathrm{Im}(\bar{\rho})$中一序列，$\{\bar{\rho}\{x_n\}\}\to z$，那么$\{\bar{\rho}\{x_n\}\}$就是一个柯西序列，从而$\{x_n\}$也是一个柯西序列. 因此，$\{x_n\}\to x\in M'$，但$\bar{\rho}$是连续的，所以$\{\bar{\rho}\{x_n\}\}\to\bar{\rho}(x)$，则$\bar{\rho}(x)=z$，所以$z\in\mathrm{Im}(\bar{\rho})$. 因此，$\bar{\rho}$是满射的，且$M'\approx M''$.

12.4　希尔伯特空间

我们已经讨论了距离空间上的拓扑性质，在此基础上我们可以继续研究内积空间，而不用受到维数的限制. 前面把注意力限定在实、复内积空间上，因此，F可以指$\mathbf{R}$或$\mathbf{C}$.

12.4.1　希尔伯特空间

我们知道F上的一个内积空间V，加之内积$<,>$：$V\times V\to F$是一个向量空间V. 如果$F=\mathbf{R}$，那么内积是双线性的；如果$F=\mathbf{C}$，那么内积是半双线性的.

在V上，内积诱导了由$\|v\|=\sqrt{<v,v>}$定义的范数.

我们特别要回忆一下范数的以下性质：

定理 12.4.1　1)（柯西—施瓦兹不等式）　对于$\forall u,v\in V$，$|<u,v>|\leqslant\|u\|\|v\|$，等号成立的条件是，当且仅当对于$\forall r\in F$，$u=rv$.

2)（三角不等式）　对于$u,v\in V$，$\|u+v\|\leqslant\|u\|+\|v\|$，等号成立的条件是，当且仅当对于$r\in F$，$u=rv$.

3)（平行四边形定律）　对于$\forall u,v\in V$，$\|u+v\|^2+\|u-v\|^2=2\|u\|^2+2\|v\|^2$.

如下所示，内积可以由范数重新加以定义.

定理 12.4.2　1) V是一个实内积空间，那么$<u,v>=1/4(\|u+v\|^2-\|u-v\|^2)$.

2）如果 V 是一个复内积空间，那么

$$< u,v > = 1/4(\| u+v \|^2 - \| u-v \|^2) + 1/4(\| u+\mathrm{i}v \|^2 - \| u-\mathrm{i}v \|^2).$$

在 V 上，内积也诱导了一个由 $d(u,v) = \| u-v \|$ 定义的距离. 因此，任意内积空间都是一个距离空间.

定义 12.4.1　令 V,W 是内积空间，且令 $\tau \in L(V,W)$.

1）如果对于 $\forall u,v \in V$，$< \tau(u),\tau(v) > = < u,v >$，称 τ 是一个**等距**.

2）双射等距叫做**等距同构**. 当 $\tau: V \to W$ 是一个等距同构时，我们就说 V 和 W 是等距同构的.

不难看出，等距总是单射的，但不一定要是满射的，即使 $V = W$.

定理 12.4.3　线性变换 $\tau \in L(V,W)$ 是一个等距 $\Leftrightarrow \tau$ 保持范数，即对于 $\forall v \in V$，$\| \tau(v) \| = \| v \|$.

定理 12.4.4　令 V 是一个内积空间，且令 $\tau \in L(V)$.

1）如果对于 $\forall v,w \in V$，$< \tau(v),w > = 0$，那么 $\tau = 0$.

2）如果 V 是一个复内积空间，且对于 $\forall v \in V$，$< \tau(v),v > = 0$，那么 $\tau = 0$.

3）一般而言，对于实内积空间，2）不成立.

我们知道内积空间就是距离空间.

如果(x_n)是内积空间 V 的一个向量序列，那么$(x_n) \to x \Leftrightarrow$ 当 $n \to \infty$ 时，$\| x_n - x \| \to 0$.

定理 12.4.5　令 V 是一个内积空间，那么

1）$(x_n) \to x, (y_n) \to y \Rightarrow < x_n,y_n > \to < x,y >$.

2）$(x_n) \to x \Rightarrow \| x_n \| \to \| x \|$.

定义 12.4.2　内积空间在由内积诱导的距离之下是完全的，称为**希尔伯特空间**.

例 12.4.1　希尔伯特空间的最重要例子之一就是 L^2. 这个内积由 $< x,y > = \sum_{n=1}^{\infty} x_n \overline{y_n}$ 定义. 由这个内积诱导的距离是 $d(x,y) = \| x-y \| = (\sum_{n=1}^{\infty} | x_n - y_n |^2)^{1/2}$. 它适合于前面所讨论的距离空间 L^2 的定义. 换句话说，前面定义的距离是由这个内积诱导的，且这个内积空间是完全的，所以希尔伯特空间也是完全的.

从例 12.4.1 中产生了一个问题：是否其他由距离

$$d(x,y) = \| x-y \|_p = (\sum_{n=1}^{\infty} | x_n - y_n |^p)^{1/p}$$

所给出的距离空间 $L^p (p \neq 2)$ 也是完全内积空间.

事实上，它们甚至不是内积空间. 更明确一点说，没有一个内积空间的诱导距离是由上式给定的. 为了说明这一点，注意到，由定理 12.4.1 知，任意来自内积的范数必满

足平行四边形定律

$$\|x+y\|^2+\|x-y\|^2=2\|x\|^2+2\|y\|^2;$$

但上式中的范数并不满足这个定律.

为了说明这一点,设 $x=(1,1,0\cdots)$, $y=(1,-1,0\cdots)$,那么

$$\|x+y\|_p=2, \|x-y\|_p=2; \|x\|_p=2^{\frac{1}{p}}, \|y\|_p=2^{\frac{1}{p}}.$$

因此,平行四边形定律的左边是 8,右边是 $4\cdot 2^{\frac{2}{p}}$,只有当 $p=2$ 时,右边才等于 8.

定理 12.4.6　令 V 是一内积空间,那么存在一个希尔伯特空间 H 和一个等距 τ: $V\to H$,使得 $\tau(V)$ 在 H 中是一个稠密子空间. 在同构意义下,H 是唯一的.

证明　距离空间(V,d),这里 d 由内积诱导,有唯一的完全化(V',d'),它由 V 中柯西序列的等价类组成. 如果$(x_n)\in\overline{(x_n)}\in V'$,$(y_n)\in\overline{(y_n)}\in V'$,那么规定

$$\overline{(x_n)}+\overline{(y_n)}=\overline{(x_n)+(y_n)}, r\overline{(x_n)}=\overline{(rx_n)},\text{且} <\overline{(x_n)},\overline{(y_n)}>=\lim_{n\to\infty}<x_n,y_n>.$$

不难看出,由于(x_n),(y_n) 是柯西序列,所以(x_n+y_n)和(rx_n)也是柯西序列.

另外,这些定义是唯一确定的,即,它们与每个等价类中代表的选取无关.

例如,如果$(\hat{x}_n)\in\overline{(x_n)}$,那么$\lim\limits_{n\to\infty}\|x_n-\hat{x},y_n\|=0$,所以

$$|<x_n,y_n>-<\hat{x}_n,y_n>|=|x_n-\hat{x}_n,y_n|\leqslant\|x_n-\hat{x}_n\|\|y_n\|\to 0.$$

(柯西序列(y_n)是有界序列.)因此,

$$<\overline{(x_n)},\overline{(y_n)}>=\lim_{n\to\infty}<x_n,y_n>=\lim_{n\to\infty}<\hat{x}_n,y_n>=<\overline{(\hat{x}_n)},\overline{(y_n)}>.$$

则 V' 是一个内积空间,而且 V' 上的内积空间诱导了距离 d'. 因为

$$\begin{aligned}<\overline{(x_n-y_n)},\overline{(x_n-y_n)}>&=\lim_{n\to\infty}<x_n-y_n,x_n-y_n>\\&=\lim_{n\to\infty}d(x_n,y_n)^2=d'((x_n),(y_n))^2,\end{aligned}$$

所以距离空间等距 τ: $V\to V'$ 是内积空间的一个等距. 因为

$$<\tau(x),\tau(y)>=d'(\tau(x),\tau(y))^2=d(x,y)^2=<x,y>,$$

从而 V' 是一个完全内积空间,$\tau(V)$ 是与 V 等距同构的 V' 的一个稠密子空间.

下一个结论与内积空间的子空间有关.

定理 12.4.7　1) 内积空间的任意完全子空间是闭的.

2) 希尔伯特空间的子空间仍是希尔伯特空间,当且仅当这个子空间是闭的.

3) 内积空间的任意有限维子空间是闭的和完全的.

证明　论断 1) 和论断 2) 可以从定理 12.2.3 中得出.

3) 设(x_n)是 S 中一个序列,$(x_n)\to x$,且 $x\notin S$. 令 $B=\{b_1,\cdots,b_m\}$ 是 S 的一个正

交哈梅尔基. S 中傅立叶展开 $s=\sum_{i=1}^{m}<x,b_i>b_i$ 有性质 $x-s\neq 0$,但是

$$<x-s,b_j>=<x,b_j>-<s,b_j>=0.$$

如果 $y=x-s, y_n=x_n-s\in S$,那么 S 中序列(y_n)收敛于正交于 S 的向量 y. 这不可能,因为 $y_n\perp y$,则 $\|y_n-y\|^2=\|y_n\|^2+\|y\|^2\geqslant\|y\|^2$ 不趋向于0,从而 S 是闭的.

为了说明内积空间上的任意有限维子空间 S 是完全子空间,我们把 S 嵌入它的完全化 S' 中,那么 S 是完全内积空间 S' 的一个有限维子空间,同样,它是一个闭子空间. 但是在 S' 中,S 是稠密的,所以 $S=S'$,这说明 S 是完全的.

12.4.2 无穷级数

由于可以在内积空间上进行向量加法和序列的收敛,所以我们可以定义无穷和或无穷级数的概念.

定义 12.4.3 令 V 是一个内积空间,且令(x_k)是 V 中一个序列,这个序列的第 n 个部分和是 $s_n=x_1+\cdots+x_n$. 如果部分和的序列(s_n)收敛于向量 $s\in V$,那么当 $n\to\infty$ 时,$\|s_n-s\|\to 0$,就说级数 $\sum x_n$ **收敛**于 s,记作 $\sum_{n=1}^{\infty}x_n=s$.

定义 12.4.4 序列 $\sum x_k$ 叫做**绝对收敛**,如果级数 $\sum_{n=1}^{\infty}\|x_k\|$ 收敛.

下一个定理给出了收敛和绝对收敛之间的重要关系.

定理 12.4.8 令 V 是一内积空间,那么 V 是一个完全空间,当且仅当级数的绝对收敛可推出级数收敛.

证明 设 V 是完全的,且 $\sum\|x_k\|<\infty$,那么部分和的序列 s_n 是一个柯西序列,因为如果 $n>m$,那么

$$\|s_n-s_m\|=\left\|\sum_{k=m+1}^{n}x_k\right\|\leqslant\sum_{k=m+1}^{n}\|x_k\|\to 0,$$

所以序列(s_n)收敛,即级数 $\sum x_k$ 收敛.

反之,绝对收敛一定收敛,且设(x_n)是 V 中的一个柯西序列. 我们要说明这个序列收敛. 由于(x_n)是一个柯西序列,所以对于 $\forall k>0$,存在 N_k,由 $i,j\geqslant N_k$ 可得 $\|x_i-x_j\|<\frac{1}{2^k}$. 显然,我们可以选出 $N_1<N_2<\cdots$,使得 $\|x_{N_{k+1}}-x_{N_k}\|<\frac{1}{2^k}$. 所以 $\sum_{k=1}^{\infty}\|x_{N_{k+1}}-x_{N_k}\|\leqslant\sum_{k=1}^{\infty}\frac{1}{2^k}<\infty$.

因此，根据假设，级数 $\sum_{k=1}^{\infty}(x_{N_{k+1}}-x_{N_k})$ 收敛. 但这是一个压缩的级数，其第 n 个部分和是 $x_{N_{n+1}}-x_{N_1}$，所以子序列 (x_{N_k}) 是收敛的. 由于含有收敛子序列的任意柯西序列必本身也收敛，所以序列 (x_k) 是收敛的，从而 V 是完全的.

12.4.3　最佳逼近

设 V 是一内积空间，S 是 V 的一个子集. 对于 $\forall x\in V$，能够在 S 中找到一个向量，在由内积诱导的距离中，这个向量与 x 最接近，这是很重要的，这样的一个向量应该存在. 这就是 V 的逼近问题.

设 $x\in V$，令 $\delta=\inf_{s\in S}\|x-s\|$，那么存在序列 s_n，使得 $\delta_n=\inf_{s\in S}\|x-s_n\|\to\delta$.

我们看一下从这个序列中可以得出什么结论. 首先，如果我们令 $y_k=x-s_k$，那么根据平行四边形定律，

$$\|y_k+y_j\|^2+\|y_k-y_j\|^2=2(\|y_k\|^2+\|y_j\|^2),$$

或

$$\|y_k-y_j\|^2=2(\|y_k\|^2+\|y_j\|^2)-4\left\|\frac{y_k+y_j}{2}\right\|^2. \tag{12.2}$$

如果集合 S 是**凸集**，即，如果对于 $\forall 0\leqslant r\leqslant 1, x,y\in S\Rightarrow rx+(1-r)y\in S$（即在其任意两点之间，$S$ 都含有一条线段），那么 $(s_k+s_j)/2\in S$，所以

$$\left\|\frac{y_k+y_j}{2}\right\|^2=\left\|x-\frac{s_k+s_j}{2}\right\|\geqslant\delta.$$

因此由(12.2)式，当 $k,j\to\infty$ 时，$\|y_k-y_j\|^2\leqslant 2(\|y_k\|^2+\|y_j\|^2)-4\delta^2\to 0$. 从而，如果 S 是凸集，那么 $(y_n)=(x-s_n)$ 是一个柯西序列，因此，(s_n) 也是一个柯西序列. 如果 S 也是完全的，柯西序列 (s_n) 收敛于向量 $\hat{s}\in S$，由范数的连续性必有，$\|x-\hat{s}\|=\delta$.

定理 12.4.9　设 V 是一内积空间，S 是 V 的完全凸子集，那么对于 $\forall x\in V$，存在唯一的 $\hat{s}\in S$，使得 $\|x-\hat{s}\|=\inf_{s\in S}\|x-s\|$. 向量 $\hat{s}$ 叫做 S 中到 x 的**最佳逼近**.

证明　只要证唯一性. 设 $\|x-\hat{s}\|=\delta=\|x-s'\|$. 根据平行四边形定律，

$$\begin{aligned}\|\hat{s}-s'\|^2&=\|(x-s')-(x-\hat{s})\|^2\leqslant 2\|x-\hat{s}\|^2+2\|x-s'\|^2-\|2x-\hat{s}-s'\|^2\\&=2\|x-\hat{s}\|^2+2\|x-s'\|^2-4\left\|x-\frac{\hat{s}+s'}{2}\right\|^2\leqslant 2\delta^2+2\delta^2-4\delta^2=0.\end{aligned}$$

所以 $\hat{s}=s'$.

由于内积空间 V 的任意子空间 S 都是凸的，所以我们可以把定理 12.4.9 应用到完全子空间上. 同时，在这一情形下，我们还有以下

定理 12.4.10　令 V 是一内积空间，S 是 V 的完全子空间，那么对于 $\forall x \in V$，S 中到 x 的最佳逼近是使 $x-s' \perp S$ 的唯一的向量 $s' \in S$.

证明　设 $x-s' \perp S$，其中 $s' \in S$，那么对于 $\forall s \in S$，$x-s' \perp s-s'$，所以

$$\| x-s \|^2 = \| x-s' \|^2 + \| s'-s \|^2 \geqslant \| x-s' \|^2.$$

因此，$s' = \hat{s}$ 是 S 中到 x 的最佳逼近. 现在我们只要说明 $x-\hat{s} \perp S$，这里 $\hat{s}$ 是 S 中到 x 的最佳逼近. 对于 $\forall s \in S$，根据完全平方的计算，

$$\begin{aligned}
\| x-rs \|^2 &= \langle x-rs, x-rs \rangle \\
&= \| x \|^2 - \bar{r}\langle x,s \rangle - r\langle s,x \rangle + r\bar{r}\| s \|^2 \\
&= \| x \|^2 + \| s \|^2 \left(r\bar{r} - \bar{r}\frac{\langle x,s \rangle}{\| s \|^2} - r\overline{\frac{\langle x,s \rangle}{\| s \|^2}} \right) \\
&= \| x \|^2 + \| s \|^2 \left(r - \frac{\langle x,s \rangle}{\| s \|^2} \right)\left(\bar{r} - \overline{\frac{\langle x,s \rangle}{\| s \|^2}} \right) - \frac{|\langle x,s \rangle|^2}{\| s \|} \\
&= \| x \|^2 + \| s \|^2 \left| r - \frac{\langle x,s \rangle}{\| s \|^2} \right|^2 - \frac{|\langle x,s \rangle|^2}{\| s \|}.
\end{aligned}$$

当 $r = r_0 = \dfrac{\langle x,s \rangle}{\| s \|^2}$ 时，最后一个表达式最小，此时 $\| x - r_0 s \|^2 \leqslant \| x \|^2 - \dfrac{|\langle x,s \rangle|^2}{\| s \|}$.

用 $x-\hat{s}$ 代替 x，得

$$\| x - \hat{s} - r_0 s \|^2 \leqslant \| x - \hat{s} \|^2 - \frac{|\langle x-\hat{s},s \rangle|^2}{\| s \|} \leqslant \delta - \frac{|\langle x-\hat{s},s \rangle|^2}{\| s \|}.$$

但 $\hat{s} + r_0 s \in S$，所以左边至少为 δ，则

$$\frac{|\langle x-\hat{s},s \rangle|^2}{\| s \|} = 0.$$

或等价地，$\langle x-\hat{s},s \rangle = 0$. 因此，$x-\hat{s} \perp S$.

根据此定理，如果 S 是内积空间 V 的一个完全子空间，那么对于 $\forall x \in V$，我们可以记为 $x = \hat{s} + (x-\hat{s})$，其中 $\hat{s} \in S$，$x-\hat{s} \in S^{\perp}$. 因此 $V = S + S^{\perp}$.

由于 $S \cap S^{\perp} = \{0\}$，所以 $V = S \odot S^{\perp}$. 这是任意内积空间上的射影定理.

定理 12.4.11(射影定理)　若 S 是内积空间 V 的一完全子空间，则 $V = S \odot S^{\perp}$. 特别地，如果 S 是希尔伯特空间 H 的一闭子空间，那么 $H = S \odot S^{\perp}$.

定理 12.4.12　令 S，T 和 T' 是内积空间 V 的子空间.

1）如果 $V = S \odot T$，那么 $T = S^{\perp}$.

2）如果 $S \oplus T = S \oplus T'$，那么 $T = T'$.

证明　1）如果 $V = S \odot T$，那么根据正交直和的定义，$T \subset S^{\perp}$. 另一方面，如果 $z \in$

$S^{\perp}$,那么对于 $\forall s \in S$ 和 $t \in T, z = s + t$. 因此,

$$0 = \langle z, s \rangle = \langle s, s \rangle + \langle t, s \rangle = \langle s, s \rangle.$$

所以 $s = 0$,则 $z = t \in T$,从而 $S^{\perp} \subset T$.

论断 2) 可以从 1) 得出.

我们用 cspan(S) 来表示向量集 S 生成的闭包.

定理 12.4.13 令 H 是一个希尔伯特空间.

1) 如果 A 是 H 的一个子集,那么 $\mathrm{cspan}(A) = A^{\perp\perp}$.

2) 如果 S 是 H 的一个子空间,那么 $cl(S) = S^{\perp\perp}$.

3) 如果 K 是 H 的一个闭子空间,那么 $K = K^{\perp\perp}$.

定理 12.4.14 如果 A 是希尔伯特空间 H 的一个子集,那么 cspan(A) 在 H 中是稠密的,当且仅当 $A^{\perp} = \{0\}$.

证明 和上面的证明一样,$H = \mathrm{cspan}(A) \odot A^{\perp}$,所以 $A^{\perp} = \{0\}$,当且仅当 $H = \mathrm{cspan}(A)$.

12.5 傅立叶级数

12.5.1 希尔伯特基

定义 12.5.1 希尔伯特空间 H 中的极大规范正交集叫做 H 的**希尔伯特基**.

可以用佐恩引理来说明任意非平凡希尔伯特空间都有一个希尔伯特基. 希尔伯特基和哈梅尔基(极大线性无关集)的概念完全不同. 后面我们将证明一个希尔伯特空间上任意两个希尔伯特基都有相同的维数.

由于规范正交集 Ω 是极大的,当且仅当 $\Omega^{\perp} = \{0\}$,所以根据定理 12.4.14,我们就有了希尔伯特基的以下性质:

定理 12.5.1 令 Ω 是希尔伯特空间 H 的一个规范正交子集,下列各条等价:

1) Ω 是一个希尔伯特基.

2) $\Omega^{\perp} = \{0\}$.

3) Ω 是 H 的**全子集**,即 $\mathrm{cspan}(\Omega) = H$.

论断 3) 表明,希尔伯特空间的子集是一个希尔伯特基,当且仅当这一子集是一个全规范正交集.

例 12.5.1 令 V 是一个内积空间,A, B 是 V 的子集,则

1) $A \subset B \Rightarrow B^{\perp} \subset A^{\perp}$.

2) $A^{\perp}$ 是 V 的一闭子空间.

3) $[\mathrm{cspan}(A)]^{\perp} = A^{\perp}$.

4）希尔伯特空间 H 的一子空间 S 是闭的 $\Leftrightarrow S = S^{\perp\perp}$.

5）令 V 是 L^2 的一个子空间，这里 L^2 由全体实数列构成，且有性质：每个序列只含有有限个非零项. 因此，V 是一个内积空间. 设 W 是 V 的一个子空间，由 V 中全体序列 $x = (x_n)$ 构成，有性质：$\sum x_n/n = 0$，则 W 是闭的，但 $W^{\perp\perp} \neq W$.

12.5.2 傅立叶展开(有限维)

现在，我们更详细地讨论一下最佳逼近的问题. 我们的目标是从希尔伯特空间 H 的一闭子空间 S 中，找出任意向量 x 的最佳逼近的一种清晰的表达式. 为了方便，我们会分三种情况讨论 ——S 是有限维的、可数无限维的、不可数维的.

设 $\Omega = \{u_1, \cdots, u_n\}$ 是希尔伯特空间 H 的一规范正交集. 关于 Ω 的 $\forall x \in H$ 的傅立叶展开由 $\hat{x} = \sum_{k=1}^{n} < x, u_k > u_k$ 给出. 这里 $< x, u_k >$ 是 x 关于 u_k 的傅立叶系数，而 $< x - \hat{x}, u_k > = < x, u_k > - < \hat{x}, u_k > = 0$，所以 $x - \hat{x} \perp \operatorname{span}(\Omega)$.

因此，根据定理 12.4.10，在 $\operatorname{span}(\Omega)$ 中，傅立叶展开 $\hat{x}$ 是到 x 的最佳逼近. 而且，由于 $x - \hat{x} \perp \hat{x}$，所以 $\| \hat{x} \|^2 = \| x \|^2 - \| x - \hat{x} \|^2 \leqslant \| x \|^2$，从而 $\| \hat{x} \| \leqslant \| x \|$ 成为等式当且仅当 $x = \hat{x}$，而这当且仅当 $x \in \operatorname{cspan}(\Omega)$ 时才会成立.

定理 12.5.2 设 $\Omega = \{u_1, \cdots, u_n\}$ 是希尔伯特空间 H 的一有限规范正交集. 对于 $\forall x \in H$，x 的傅立叶展开 $\hat{x}$ 是 $\operatorname{span}(\Omega)$ 中到 x 的最佳逼近. 我们也有贝塞尔不等式 $\| \hat{x} \| \leqslant \| x \|$. 或等价地

$$\sum_{k=1}^{n} |< x, u_k >|^2 \leqslant \| x \|^2,$$

等式成立 $\Leftrightarrow x \in \operatorname{span}(\Omega)$.

12.5.3 傅立叶展开(可数)

在可数无限维的情形下会涉及无穷和，这样就会出现收敛的问题.

定理 12.5.3 设 $\Omega = \{u_1, u_2, \cdots\}$ 是希尔伯特空间 H 中的一可数无限规范正交集. 级数 $\sum_{k=1}^{\infty} r_k u_k$ 在 H 中收敛，当且仅当级数 $\sum_{k=1}^{\infty} | r_k |^2$ 在 $\mathbf{R}$ 中收敛. 如果这些级数收敛，那么它们就无条件收敛.

最后，如果级数 $\sum_{k=1}^{\infty} r_k u_k$ 收敛，那么 $\| \sum_{k=1}^{\infty} r_k u_k \|^2 = \sum_{k=1}^{\infty} | r_k |^2$.

证明 用 s_n 表示第一个级数中的部分和，p_n 表示第二个级数中的部分和，那么当 $m \leqslant n$ 时有

$$\| s_n - s_m \|^2 = \| \sum_{k=m+1}^{\infty} r_k u_k \|^2 = \sum_{k=m+1}^{\infty} | r_k |^2 = | p_n - p_m |.$$

因此,(s_n) 是 H 中的一个柯西序列,当且仅当(p_n) 是 $\mathbf{R}$ 中的柯西序列. 由于 H 和 $\mathbf{R}$ 都是完全的,所以(s_n) 收敛,当且仅当(p_n) 收敛.

若级数 $\sum_{k=1}^{\infty} | r_k |$ 收敛,则它就绝对收敛,从而无条件收敛.

但是如果 $\sum_{k=1}^{\infty} | r_k |$ 无条件收敛,那么级数 $\sum_{k=1}^{\infty} r_k u_k$ 也无条件收敛.

现在设 $\Omega = \{u_1, u_2, \cdots\}$ 是 H 中的一个可数无限规范正交集. 向量 $x \in H$ 的傅立叶展开定义为和 $\hat{x} = \sum_{k=1}^{\infty} < x, u_k > u_k$. 对于 $\forall n > 0$,由 $\| \hat{x} \| \leqslant \| x \|$ 得, $\sum_{k=1}^{n} |< x, u_k >|^2 \leqslant \| x \|^2$,所以

$$\sum_{k=1}^{\infty} |< x, u_k >|^2 \leqslant \| x \|^2.$$

这说明上式左边的级数收敛. 因此,根据定理 12.5.3,傅立叶级数

$$\hat{x} = \sum_{k=1}^{\infty} < x, u_k > u_k$$

无条件收敛. 而且,由于内积是连续的,$< x - \hat{x}, u_k > = < x, u_k > - < \hat{x}, u_k > = 0$,所以 $x - \hat{x} \in [\text{span}(\Omega)]^{\perp} = [\text{cspan}(\Omega)]^{\perp}$. 因此,$\hat{x}$ 是 cspan(Ω) 中到 x 的最佳逼近.

最后,由于 $x - \hat{x} \perp \hat{x}$,所以又有 $\| \hat{x} \|^2 = \| x \|^2 - \| x - \hat{x} \|^2 \leqslant \| x \|^2$,从而 $\| \hat{x} \| \leqslant \| x \|$ 等号成立的充分必要条件是 $x = \hat{x}$,而这当且仅当 $x \in \text{cspan}(\Omega)$ 时才成立.

因此,以下定理与定理 12.5.2 类似.

定理 12.5.4 令 $\Omega = \{u_1, \cdots, u_n\}$ 是希尔伯特空间 H 中的一可数无限规范正交集. 对于 $\forall x \in H$,x 的傅立叶展开

$$\hat{x} = \sum_{k=1}^{\infty} < x, u_k > u_k$$

无条件收敛,是 cspan(Ω) 中到 x 的最佳逼近. 我们也有**贝塞尔不等式** $\| \hat{x} \| \leqslant \| x \|$. 或等价地,$\sum_{k=1}^{\infty} |< x, u_k >|^2 \leqslant \| x \|^2$,等式成立的充分必要条件是 $x \in \text{cspan}(\Omega)$.

12.5.4 傅立叶级数

为了讨论任意规范正交集 $\Omega = \{u_k \mid k \in K\}$ 的情况,我们首先定义并讨论一个任意项数的和的概念.

定义 12.5.2　令 $\Xi=\{x_k \mid k\in K\}$ 是内积空间 V 中的一任意向量族. 如果对于 $\forall\varepsilon>0$,都存在一有限集 $S\subset K$,使当 $T\supset S$,T 有限时得,$\|\sum\limits_{k\in K}x_k-x\|\leqslant\varepsilon$,那么和 $\sum\limits_{k\in K}x_k$ 叫做**收敛**于向量 $x\in V$,记为 $x=\sum\limits_{k\in K}x_k$.

对于熟悉网收敛定义的读者来说,K 的全体有限子集所成的集合 $P_0(K)$ 是在包含关系下的一有限集,函数 $S\to\sum\limits_{k\in K}x_k$ 是 H 中的一个网. $x=\sum\limits_{k\in K}x_k$ 的收敛就是这个网的收敛. 无论哪种情况,我们都称上面的定义为收敛的网定义.

对于以下任意和的网收敛的基本性质,并不难验证.

定理 12.5.5　令 $\widetilde{H}=\{x_k \mid k\in K\}$ 是内积空间 V 中一任意向量族,若 $\sum\limits_{k\in K}x_k=x$, $\sum\limits_{k\in K}y_k=y$,那么

1) 对于 $\forall r\in F$, $\sum\limits_{k\in K}rx_k=rx$.

2) $\sum\limits_{k\in K}(x_k+y_k)=x+y$.

3) $\sum\limits_{K}<x_k,y>=<x,y>$, $\sum\limits_{K}<y,x_k>=<y,x>$.

定理 12.5.6　令 $\widetilde{H}=\{x_k \mid k\in K\}$ 是内积空间 V 中一任意向量族.

1) 如果和 $\sum\limits_{k\in K}x_k$ 收敛,那么对于 $\forall\varepsilon>0$,存在一有限集 $I\subset K$,使得 $J\cap I=\varnothing$,若 J 是有限的,则 $\|\sum\limits_{k\in J}x_k\|\leqslant\varepsilon$.

2) 如果 V 是一希尔伯特空间,那么论断 1) 的逆命题也成立.

证明　1) 给定 $\varepsilon>0$,设 $S\subset K$,S 有限,且当 $T\supset S$,T 有限时,

$$\|\sum_{k\in T}x_k-x\|\leqslant\frac{\varepsilon}{2}.$$

如果 $J\cap S=\varnothing$,J 是有限的,那么

$$\begin{aligned}\|\sum_{J}x_k\|&=\|(\sum_{J}x_k+\sum_{S}x_k-x)-(\sum_{S}x_k-x)\|\\&\leqslant\|\sum_{J\cup S}x_k-x\|+\|\sum_{S}x_k-x\|\leqslant\frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon.\end{aligned}$$

2) 对于 $\forall n>0$,设 $I_n\subset K$ 是有限集,使得 $J\cap I_n=\varnothing$,当 J 有限时,则 $\|\sum\limits_{j\in J}x_j\|\leqslant\frac{1}{n}$,且设 $y_n=\sum\limits_{k\in I_n}x_k$,那么$(y_n)$是一个柯西序列,因为 $\|y_n-y_m\|=\|\sum\limits_{I_n}x_k-\sum\limits_{I_m}x_k\|=\|\sum\limits_{I_n-I_m}x_k-\sum\limits_{I_m-I_n}x_k\|\leqslant\|\sum\limits_{I_n-I_m}x_k\|+\|\sum\limits_{I_m-I_n}x_k\|\leqslant\frac{1}{m}+\frac{1}{n}\to0$,由于设 V 为一个完全空间,

所以 $(y_n) \to y$.

现在,给定 $\varepsilon > 0$,存在一个 N,使得 $n \geqslant N \Rightarrow \| y_n - y \| = \| \sum_{I_n} x_k - y \| \leqslant \frac{\varepsilon}{2}$.

规定 $n = \max\{N, 2/\varepsilon\}$,得 $T \supset I_n$,T 有限,则

$$\| \sum_T x_k - y \| = \| \sum_{I_n} x_k - y + \sum_{T-I_n} x_k \| \leqslant \| \sum_{I_n} x_k - y \| + \| \sum_{T-I_n} x_k \| \leqslant \frac{\varepsilon}{2} + \frac{1}{n} \leqslant \varepsilon.$$

所以 $\sum_{k \in K} x_k$ 收敛于 y.

定理 12.5.7 令 $\widetilde{H} = \{x_k \mid k \in K\}$ 是内积空间 V 的任一向量族.若和 $\sum_{k \in K} x_k$ 收敛,那么最多有可数个项 x_k 不为 0.

证明 根据定理 12.5.6,对于 $\forall n > 0$,我们可以设 $I_n \subset K$,I_n 是有限的,使得 $J \cap I_n = \varnothing$,J 是有限的,则 $\| \sum_{j \in J} x_j \| \leqslant \frac{1}{n}$.设 $I = \bigcup_n I_n$,那么 I 是可数的,且对于 $\forall n, k \notin I$,当 $\{k\} \cap I_n = \varnothing$ 时,对于 $\forall n$,$\| x_k \| \leqslant \frac{1}{n}$,则 $x_k = 0$.

定理 12.5.8 令 $\widetilde{H} = \{u_k \mid k \in K\}$ 是希尔伯特空间 H 中一任意规范正交向量族.两级数 $\sum_{k \in K} r_k u_k$ 和 $\sum_{k \in K} | r_k |^2$ 一起收敛或发散,如果这些级数收敛,那么

$$\left\| \sum_{k \in K} r_k u_k \right\|^2 = \sum_{k \in K} | r_k |^2.$$

证明 第一个级数收敛 $\Leftrightarrow$ 对于 $\forall \varepsilon > 0$,存在一有限集 $I \subset K$,使得 $J \cap I = \varnothing$,J 是有限的,则 $\left\| \sum_{k \in J} r_k u_k \right\|^2 \leqslant \varepsilon^2$.等价地,$J \cap I = \varnothing$,$J$ 是有限时,有 $\sum_{k \in J} \| r_k \|^2 \leqslant \varepsilon^2$.而这恰指的是第二个级数收敛.

定理 12.5.9 令 $\{r_k \mid k \in K\}$ 是一非负实数集,如果上述表达之一是有限的,那么

$$\sum_{k \in K} r_k = \sup_{\substack{J \text{有限} \\ J \subset K}} \sum_{k \in J} r_k. \tag{12.3}$$

证明 设 $\sup_{\substack{J \text{有限} \\ J \subset K}} \sum_{k \in J} r_k = R < \infty$,则 $\forall \varepsilon > 0$,存在 $S \subset K$,使 $\mathbf{R} \geqslant \sum_{k \in S} r_k \geqslant \mathbf{R} - \varepsilon$.

如果 $T \subset K$ 是一有限集,其中 $T \supset S$,那么由 $r_k \geqslant 0$,$\mathbf{R} \geqslant \sum_{k \in T} r_k \geqslant \sum_{k \in S} r_k \geqslant \mathbf{R} - \varepsilon$,可得 $\| \mathbf{R} - \sum_{k \in T} r_k \| \leqslant \varepsilon$.这说明 $\sum r_k$ 收敛于 $\mathbf{R}$.如果 $\sum_{k \in K} r_k$ 收敛,那么(12.3) 式右边的上确界是有限的,所以(12.3) 式成立.

对于可数无限级数的收敛,我们有两个定义 —— 网收敛的定义和包含部分和的极限的传统定义.分别用 $\sum_{k \in \mathbf{N}^*} x_k$ 和 $\sum_{k=1}^{\infty} x_k$ 来表示网收敛的定义和部分和的定义.

定理 12.5.10 设 H 是一个希尔伯特空间. 对于 $\forall k, x_k \in H$, 下列论断等价:

1) $\sum\limits_{k \in \mathbf{N}^*} x_k$ 收敛于 x(网收敛的定义); 2) $\sum\limits_{k=1}^{\infty} x_k$ 无条件收敛于 x.

证明 1)⇒2) 设 π 是 $\mathbf{N}^+$ 的任意排列. 给定 $\forall \varepsilon > 0$, 存在一有限集 $S \subset \mathbf{N}^+$, 使得 $T \supset S$, T 是有限时, 有 $\| \sum\limits_{k \in T} x_k - x \| \leqslant \varepsilon$. 我们用 I_n 表示整数集 $\{1, \cdots, n\}$, 同时选出一正整数 n, 使得 $\pi(I_n) \supset S$, 则当 $m \geqslant n$ 时, 有 $\pi(I_m) \supset \pi(I_n) \supset S$, 则

$$\| \sum_{k=1}^{m} x_{\pi(k)} - x \| = \| \sum_{k \in \pi(I_m)} x_k - x \| \leqslant \varepsilon.$$

2)⇒1) 因为 1) 中的级数不收敛, 所以存在 $\varepsilon > 0$, 使得对于任意有限子集 $I \subset \mathbf{N}^+$, 存在一有限子集 J, 使得 $J \cap I = \varnothing$, $\| \sum\limits_{k \in J} x_k \| > \varepsilon$, 则存在 $\mathbf{N}^+$ 的互不相交有限子集的可数无限序列 J_n, 且

$$\max(J_n) = M_n < m_{n+1} = \min(J_{n+1}); \| \sum_{k \in J_n} x_k \| > \varepsilon.$$

现在, 我们选取任意排列 π: $\mathbf{N}^+ \to \mathbf{N}^+$, 它有以下性质:

1) $\pi([m_n, M_n]) \subset [m_n, M_n]$.

2) 若 $J_n = \{j_{n,1}, \cdots, j_{n,\mu}\}$, 则 $\pi(m_n) = j_{n,1}$, $\pi(m_n + 1) = j_{n,2}$, $\cdots$, $\pi(m_n + u_n - 1) = j_{n,\mu}$.

对于这样的任意排列 π, 我们有 $\| \sum\limits_{k=m_n}^{m_n+u_n-1} x_{\pi(k)} \| = \| \sum\limits_{k \in J_n} x_k \| > \varepsilon$.

这说明级数的部分和的序列 $\sum\limits_{k=1}^{\infty} x_{\pi(k)}$ 不是柯西序列, 所以这个级数不收敛, 这与 2) 矛盾; 同时说明, 由论断 2) 至少可得论断 1) 收敛. 但如果论断 1) 收敛于 $y \in H$, 那么由 1) 可得 2), 又因为无条件极限是唯一的, 所以 $y = x$. 因此, 由 2) 可得 1).

设 $\Omega = \{u_k \mid k \in K\}$ 是希尔伯特空间 H 中一任意规范正交集. 给定 $\forall x \in H$, 由定理 12.5.2 得

$$\sup_{\substack{J \text{ 有限} \\ J \subset K}} \sum_{k \in J} |< x, u_k >|^2 \leqslant \| x \|^2.$$

所以由定理 12.5.9 知和 $\sum\limits_{k \in K} |< x, u_k >|^2$ 收敛. 因此, 根据定理 12.5.8, x 的傅立叶级数 $\hat{x} = \sum\limits_{k \in K} < x, u_k > u_k$ 也收敛, 且 $\| \hat{x} \|^2 = \sum\limits_{k \in K} |< x, u_k >|^2$. 根据定理 12.5.7, $\hat{x}$ 是 $< x, u_k > u_k$ 的项的一个可数无限和, 所以 $\hat{x}$ 在 $\operatorname{cspan}(\widetilde{H})$ 中, 且

$$< x - \hat{x}, u_k > = < x, u_k > - < \hat{x}, u_k > = 0.$$

所以 $x-\hat{x}\in[\mathrm{span}(\Omega)]^{\perp}=[\mathrm{cspan}(\Omega)]^{\perp}$. 因此,$\hat{x}$ 是 $\mathrm{cspan}(\Omega)$ 中到 x 的最佳逼近.

最后,由于 $x-\hat{x}\perp\hat{x}$,所以又有 $\|\hat{x}\|^2=\|x\|^2-\|x-\hat{x}\|^2\leqslant\|x\|^2$,所以 $\|\hat{x}\|\leqslant\|x\|$. 等式成立的条件是 $x=\hat{x}$,而这当且仅当 $x\in\mathrm{cspan}(\Omega)$ 时才会成立.

定理 12.5.11　令 $\Omega=\{u_k\mid k\in K\}$ 是希尔伯特空间 H 中一规范正交向量族. 对于 $\forall x\in H$,x 的傅立叶展开 $\hat{x}=\sum\limits_{k\in K}<x,u_k>u_k$ 在 H 中收敛,且是 $\mathrm{cspan}(\Omega)$ 中到 x 的最佳逼近.

还有贝塞尔不等式 $\|\hat{x}\|\leqslant\|x\|$. 或等价地,$\sum\limits_{k\in K}|<x,u_k>|^2\leqslant\|\hat{x}\|^2$,等式成立的充分必要条件是 $x\in\mathrm{cspan}(\Omega)$.

12.6　希尔伯特空间的特征

12.6.1　希尔伯特基的维数

在定理 12.4.14 中,希尔伯特空间 H 中一规范正交集 $\Omega=\{u_k\mid k\in K\}$ 是一个希尔伯特基,当且仅当 $\mathrm{cspan}(\Omega)=H$.

因此,从定理 12.5.11 中可以得到希尔伯特基的以下性质:

定理 12.6.1　令 $\Omega=\{u_k\mid k\in K\}$ 是希尔伯特空间 H 中一规范正交族,下列各条等价:

1) Ω 是一个希尔伯特基(一个极大规范正交集).

2) $\Omega^{\perp}=\{0\}$.

3) $\mathrm{cspan}(\Omega)=H$.

4) 对于 $\forall x\in H$,$x=\hat{x}$.

5) 在贝塞尔不等式中,对于 $\forall x\in H$,等式成立,即对于 $\forall x\in H$,$\|x\|=\|\hat{x}\|$.

6) 对于 $\forall x,y\in H$,帕塞瓦尔恒等式 $<x,y>=<\hat{x},\hat{y}>$ 成立,即

$$<x,y>=\sum_{k\in K}<x,u_k>\overline{<y,u_k>}.$$

现在我们讨论希尔伯特空间 H 的全体希尔伯特基都有相同的基数,从而 H 的希尔伯特维数可以定义为这一基数.

定理 12.6.2　希尔伯特空间 H 的全体希尔伯特基都有相同的基数. 这个基数称为 H 的 **希尔伯特维数**. 用 $\mathrm{hdim}(H)$ 表示 H 的希尔伯特维数.

证明　如果 H 含有一个有限希尔伯特基,那么这个集合也是一个哈梅尔基,所以全体希尔伯特基的维数都为 $\dim(H)$.

下面假设 $B=\{b_k \mid k\in K\}$ 和 $X=\{c_j \mid j\in J\}$ 是 H 的无限希尔伯特基，那么对于 $\forall\, b_k, b_k=\sum_{j\in J_k}<b_k,c_j>c_j$. 这里 J_k 是一个可数集 $\{j \mid <b_k,c_j>\neq 0\}$. 而且，由于没有一个 c_j 可以与每个 b_k 都正交，所以 $|\bigcup J_k|=J$. 由于任意 J_k 都是可数的，因此

$$|J|=|\bigcup J_k|\leqslant \aleph_0 |K|=|K|.$$

由对称性得 $|K|\leqslant|J|$，由伯恩斯基定理可知 $|K|=|J|$.

定理 12.6.3　两个希尔伯特空间等距同构 $\Leftrightarrow$ 它们有相同的希尔伯特维数.

证明　设 $\mathrm{hdim}(H_1)=\mathrm{hdim}(H_2)$. 设 $\Omega_1=\{u_k \mid k\in K\}$ 是 H_1 的一个希尔伯特基，$\Omega_2=\{v_k \mid k\in K\}$ 是 H_2 的一个希尔伯特基. 我们可以如下定义映射

$$\tau:\ H_1\rightarrow H_2,\tau(\sum_{k\in K} r_k u_k)=\sum_{k\in K} r_k v_k.$$

容易验证 τ 是一个双射等距.

12.6.2　希尔伯特空间的特征

任意向量空间 V 都同构于从 B 到 F 的含有有限支集的全体函数所成的向量空间 $(F^B)_0$. 希尔伯特空间中也有一个相应的结论. 令 K 是任意非空集合，且令

$$L^2(K)=\{f:\ K\rightarrow \mathbf{C} \mid \sum_{k\in K}|f_k|^2<\infty\},$$

$L^2(K)$ 中的函数称为**平方可和函数**.

由 $<f,g>=\sum_{k\in K}(f_k)\overline{g}_k$，我们可以在 $L^2(K)$ 上定义内积.

$L^2(K)$ 是希尔伯特空间的证明与 $L^2=L^2(\mathbf{N})$ 是希尔伯特空间的证明类似. 由 $\delta_k(j)=\delta_{k,j}$ 定义的 $\delta_k\in L^2(K)$，那么集合 $\Omega=\{\delta_k \mid k\in K\}$ 为 $L^2(K)$ 的基数是 $|K|$ 的一个希尔伯特基.

因为

$$<\delta_i,\delta_j>=\sum_{k\in K}\delta_i(k)\overline{\delta_j(k)}=\delta_{i,j},$$

所以 Ω 是一个规范正交集. 而且，如果 $f\in L^2(K)$，那么对于可数个 $k\in K$，设为 $\{k_1,k_2,\cdots\}$，$f(k)\neq 0$.

如果由 $f'=\sum_{i=1}^{\infty} f(k_i)\delta_{k_i}$ 定义 f'，那么 $f'\in \mathrm{cspan}(\Omega)$，且对于 $\forall\, j\in K, f'(j)=f(j)$，则 $f=f'$，从而 $L^2(K)=\mathrm{cspan}(\Omega)$，所以 Ω 是一个完全规范正交集，即 Ω 是 $L^2(K)$ 的一个希尔伯特基. 设 H 是一个含有希尔伯特基 $B=\{u_k \mid k\in K\}$ 的希尔伯特空间.

定义映射 $\varphi:\ H\rightarrow l^2(K)$. 由于 B 是一个希尔伯特基，所以 $\forall\, x\in H$ 都有形式 $x=$

$\sum_{k\in K} < x, u_k > u_k$. 由于级数收敛，所以由定理 12.5.8 知级数 $\sum_{k\in K} |< x, u_k >|^2$ 收敛. 因此，由定理 12.5.8 知级数 $\varphi(x)$ 收敛，规定 $\varphi(x) = \sum_{k\in K} |< x, u_k >| \delta_k$. 从定理 12.4.6 得出 φ 是一个线性映射，而且它也是双射，

因为 $\varphi(u_k) = \delta_k$，所以 φ 取 H 的希尔伯特基 B 到 $L^2(K)$ 的希尔伯特基 Ω. 同样，

$$\| \varphi(x) \|^2 = < \varphi(x), \varphi(x) > = \sum_{k\in K} |< x, u_k >|^2 = \| \sum_{k\in K} < x, u_k > u_k \|^2 = \| x \|^2,$$

所以 φ 是一个等距同构. 我们已经证明了以下

定理 12.6.4　如果 H 是一个维数为 k 的希尔伯特空间，K 是基数为 k 的任意集合，那么 H 与 $L^2(K)$ 是等距同构的.

12.6.3　里斯表示定理

如果 $y \in H$，那么函数 $f_y(x) = < x, y >$ 必为 H 上一线性泛函. 特别地，由施瓦茨不等式，对于 $\forall x \in H$，$| f_y(x) | = |< x, y >| \leqslant \| x \| \| y \|$，或对于 $\forall x \neq 0$，

$$\frac{| f_y(x) |}{\| x \|} \leqslant \| y \|.$$

如果 $x = y$，那么上式等号成立，我们有

$$\sup_{x\neq 0} \frac{| f_y(x) |}{\| x \|} = \| y \|,$$

这样，我们就有了以下定义，我们把它用在希尔伯特空间之间的线性变换上(包括线性泛函的情况).

定义 12.6.1　令 τ: $H_1 \to H_2$ 是一个线性变换，如果 $\sup_{x\neq 0} \frac{\| \tau(x) \|}{\| x \|} < \infty$，那么 τ 称为是**有界的**. 如果上式上确界是有限的，我们用 $\| \tau \|$ 来表示，称为 τ 的**范数**.

当然，如果 f: $H \to F$ 是 H 上一有界线性泛函，那么

$$\| f \| = \sup_{x\neq 0} \frac{| f(x) |}{\| x \|}.$$

希尔伯特空间 H 上全体有界线性泛函所成的集合称为 H 的**连续对偶空间**，或**共轭空间**，表示为 H^*. 它与 H 的代数对偶空间不同，代数对偶空间指 H 上全体线性泛函所成的集合. 但在有限维的情形下，由于所有线性泛函都是有界的，所以这两个概念是一致的.

定理 12.6.5　令 τ: $H_1 \to H_2$ 是一个有界线性变换.

1) $\| \tau \| = \sup_{\| x \| = 1} \| \tau(x) \|$.

2) $\| \tau \| = \sup_{\| x \| \leqslant 1} \| \tau(x) \|$.

3) $\|\tau\| = \inf\{c \in \mathbf{R}\}$,对于 $\forall x \in H$, $\|\tau(x)\| \leqslant c\|x\|\}$.

定理 12.6.6 令 $\tau: H_1 \to H_2$ 是一有界线性变换,那么以下各条等价:

1) τ 是有界的.

2) τ 在任意点 $x_0 \in H$ 处是连续的.

3) τ 是连续的.

证明 设 τ 有界,那么当 $x \to x_0$ 时, $\|\tau(x) - \tau(x_0)\| = \|\tau(x - x_0)\| \leqslant \|\tau\| \ \|x - x_0\| \to 0$,所以 τ 在 x_0 处连续. 因此,由论断 1) 可得 2). 如果论断 2) 成立,那么对于 $\forall y \in H$,当 $x \to y$ 时, $\|\tau(x) - \tau(y)\| = \|\tau(x - y + x_0) - \tau(x_0)\| \to 0$. 由于 τ 在 x_0 处是连续的,并且当 $y \to x$ 时, $x - y + x_0 \to x_0$,所以 τ 在 $\forall y \in H$ 处连续,从而论断 3) 成立.

再设 3) 成立,则 τ 在 0 处连续,所以存在 $\delta > 0$,由 $\|x\| \leqslant \delta$,可得 $\|\tau(x)\| \leqslant 1$. 特别地,由 $\|x\| = \delta$ 可得 $\dfrac{\|\tau(x)\|}{\|x\|} \leqslant \dfrac{1}{\delta}$,所以

$$\|x\| = 1 \Rightarrow \|\delta x\| = \delta \Rightarrow \frac{\|\tau(\delta x)\|}{\|\delta x\|} \leqslant \frac{1}{\delta} \Rightarrow \frac{\|\tau(x)\|}{\|x\|} \leqslant \frac{1}{\delta}.$$

因此, τ 是有界的.

定理 12.6.7(里斯表示定理) 令 H 是一个希尔伯特空间. 对于 H 上任意有界线性泛函 f,都存在唯一的 $z_0 \in H$,使得对于 $\forall x \in H$, $f(x) = \langle x, z_0 \rangle$,且 $\|z_0\| = \|f\|$.

证明 如果 $f = 0$,那么我们可以取 $z_0 = 0$,所以我们假定 $f \neq 0$. 因此, $K = \mathrm{Ker}(f) \neq H$,由于 f 连续,所以 K 是闭空间. 因此, $H = K \odot K^{\perp}$.

现在,把第一同构定理应用到线性泛函 $f: H \to F$,则 $H/K \cong F$(一个向量空间). 另外,由于 $H/K \cong K^{\perp}$,所以 $K^{\perp} \cong F$. 特别地, $\dim(K^{\perp}) = 1$.

对于 $\forall z \in K^{\perp}$,由 $x \in K$,可得 $f(x) = 0 = \langle x, z \rangle$.

又由于 $\dim(K^{\perp}) = 1$,所以我们要找出一个 $0 \neq z \in K^{\perp}$,使得 $f(z) = \langle z, z \rangle$. 那么对于 $\forall r \in F$,同样, $f(rz) = rf(z) = r\langle z, z \rangle = \langle rz, z \rangle$,说明对于 $x \in K$,有 $f(x) = \langle x, z \rangle$.

但是如果 $0 \neq z \in K^{\perp}$,那么容易验证 $z_0 = \dfrac{\overline{f(z)}}{\langle z, z \rangle} z$ 有上述性质,从而 $\|z_0\| = \|f\|$.

第 13 章　向量空间的张量积

在前面几章中，我们看到了几种从一个向量空间构造新的向量空间的方法，其中最重要的两种构造是直和 $U \oplus V$ 及 U 到 V 的全体线性变换所成的集合 $\Lambda(U,V)$，本章我们会研究另一种构造——张量积.

我们首先重新定义一下大家所熟悉的外直和、双线性映射等概念，然后讨论张量积、张量积的性质、线性变换的张量积及多重张量积，我们还需要自由向量空间的概念.

13.1　自由向量空间

13.1.1　自由向量空间

定义 13.1.1　令 V 是一域. 给定任意非空集 X，我们可以在 F 上构造一个基为 X 的向量空间 $\mathcal{F}_X$，即取 $\mathcal{F}_X$ 为 X 中元素的全体形式有限线性组合所成的集合

$$\mathcal{F}_X = \{\text{有限和} \sum r_i x_i \mid x_i \in X, r_i \in F\}.$$

其运算是运用法则 $rx_i + sx_i = (r+s)x_i$；$r(sx_i) = (rs)x_i$ 组合同类项. 向量空间 $\mathcal{F}_X$ 称为 X 上的**自由向量空间**.

事实上，任意向量空间 V 在其任意基上都是自由向量空间. 因而，从某种意义上说，我们并没有介绍新内容. 但是，自由对象的概念在本书中已出现多处，比如模的相关内容，并不是所有模都是自由的. 而且，在向量空间中，我们可以从中继续讨论相关内容.

我们可以把自由向量空间 $\mathcal{F}_X$ 特征化为 X 到 F 上含有有限支集的全体函数所成的集合 $(F^X)_0$. 函数 $f: X \to F$ 的支集由 $\operatorname{supp}(f) = \{x \in X \mid f(x) \neq 0\}$ 定义. 容易看出含有有限支集的函数 $f: X \to F$，通过 $f \mapsto \sum f(x_i)x_i$ 与 X 中元素的一个有限和对应. 因此，$\mathcal{F}_X$ 的这两种构造是等价的，我们可以任取两种构造之一.

如下所示，我们可以更一般地表达自由的概念. 考虑由 $j(x) = x$ 定义的映射 $j: X \to \mathcal{F}_X$，称为从 X 到 $\mathcal{F}_X$ 的**典范单射**.

$(\mathcal{F}_X, j)$ 有一种很特殊的性质. 在图 13-1 中，如果 $f: X \to V$ 是 X 到任意向量空间 V 的任意映射，那么从 $\mathcal{F}_X$ 到 V 存在唯一的线性变换 τ，使得 $\tau \circ j = f$.

如果 $f: X \to V$，那么通过规定 $\tau(x) = f(x)$，由线性性扩充到 $\mathcal{F}_X$，我们可以定义线性变换 $\tau: \mathcal{F}_X \to V$. 这是合理的，因为 X 是 $\mathcal{F}_X$ 的一个基. 同样，τ 的唯一性可以从 X 是 $\mathcal{F}_X$ 的一个基这一事实中得出.

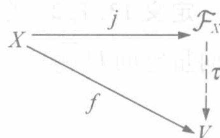


图 13－1

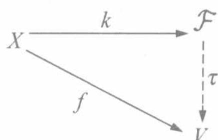
当一个图中在同一位置起始、终止的任意两条路径描述相同的函数时，比如图 13－1，我们就说这个图**可交换**.

因此，假定 $\tau \circ j = f$ 与假定图 13－1 中的图可交换是等价的. 通过典范单射 j，任意函数 $f: X \to V$ 都可以分解，这一表达也可以描述上述情形.

现在，图 13－1 的交换性和 τ 的唯一性就可以确定$(\mathcal{F}_X, j)$了.

更具体一点，我们有以下定理 —— 自由向量空间 $\mathcal{F}_X$ 的泛性质.

定理 13.1.1(自由向量空间 $\mathcal{F}_X$ 的泛性质)　令 X 是一非空集，$\mathcal{F}$ 是 F 上一向量空间，$k: X \to \mathcal{F}$ 是一个函数，$(\mathcal{F}, k)$ 有以下性质：在图 13－2 中，对于任意函数 $f: X \to V$，都存在唯一的线性变换 $\tau: \mathcal{F} \to V$，使得 $\tau \circ k = f$，即，使得图 13－2 中的图可交换，其中 V 是 F 上一个向量空间，那么 $\mathcal{F}$ 同构于自由向量空间 $\mathcal{F}_X$.



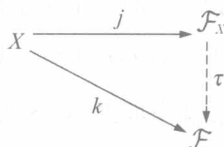
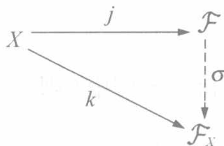


图 13－2

证明　考虑到图 13－2 中的第二张图反映出在图 13－1 中，我们可以使 $V = \mathcal{F}$. 由于图 13－2 可交换，所以 $\tau \circ j = k$.

图 13－3 中的第一张图反映出在图 13－2 中，我们可以规定 $V = \mathcal{F}_X$. 由于图 13－3 中第一张图可交换，所以 $\sigma \circ \kappa = j$. 作适当代换得 $\tau \circ \sigma \circ k = k$，$\sigma \circ \tau \circ j = j$.

但图 13－3 中的第二张交换图指出恒等变换是唯一的使得 $\iota \circ k = k$ 的线性变换，所以 $\tau \circ \sigma = \iota$. 同样，通过画相关的交换图，我们推出 $\sigma \circ \tau = \iota$. 因此，$\tau$ 是 $\mathcal{F}_X$ 到 $\mathcal{F}$ 的一个同构.



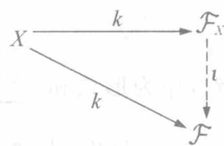


图 13－3

13.1.2　直和的特性

下面我们讨论一下外直和的构造. 我们将用三种方法来描述这种直和的特征.

定义 13.1.2 令 U,V 是同一域 F 上的向量空间. **外直和** $U\overline{\oplus}V$ 是全体有序对所成的向量空间 $U\overline{\oplus}V=\{(u,v)\mid u\in U,v\in V\}$,以及坐标式运算

$$(u,v)+(u',v')=(u+u',v+v');$$

$$r(u,v)=(ru,rv).$$

至于它的第二种性质,我们首先考虑笛卡儿积 $U\times V$,它是没有代数结构的全体有序对所成的集合 $U\times V=\{(u,v)\mid u\in U,v\in V\}$. 令$\mathcal{F}_{U\times V}$ 是 $U\times V$ 上的自由向量空间,因而,$\mathcal{F}_{U\times V}=\{$有限和$\sum r_i(u_i,v_i)\mid(u_i,v_i)\in U\times V,r_i\in F\}$. 在$\mathcal{F}_{U\times V}$ 中,不能对有序基的坐标随意操作. 比如,我们既不能用(ru,rv) 代替 $r(u,v)$,也不能用$(u+u',v+v')$ 代替$(u,v)+(u',v')$.

事实上,$\mathcal{F}_{U\times V}$ 和 $U\overline{\oplus}V$ 的区别在于,在 $U\overline{\oplus}V$ 中,对于 $\forall r\in F,u\in U$ 和 $v\in V$,有

$$r(u,v)-(ru,rv)=0;\quad (u,v)+(u',v')-(u+u',v+v')=0.$$

把 S 定义为$\mathcal{F}_{U\times V}$ 的一个子空间,对于 $\forall r\in F,u\in U$ 和 $v\in V$,S 由形式为

$$r(u,v)-(ru,rv)=0 \text{ 和 } (u,v)+(u',v')-(u+u',v+v')=0$$

的全体向量生成. 商空间$\mathcal{F}_{U\times V}/S$ 必同构于直和 $U\overline{\oplus}V$.

考虑由 $\tau(\sum r_i(u_i,v_i)+S)=\sum r_i(u_i,v_i)$ 定义的映射 τ: $\mathcal{F}_{U\times V}/S\to U\overline{\oplus}V$. 这一映射是唯一确定的,因为如果$\sum r_i(u_i,v_i)+S=\sum s_i(x_i,y_i)+S$,那么 $\sum r_i(u_i,v_i)-\sum s_i(x_i,y_i)\in S$. 但 S 中任意元都等于 $U\overline{\oplus}V$ 中的零向量,所以向量$\sum r_i(u_i,v_i)$ 和 $\sum s_i(x_i,y_i)$ 在 $U\overline{\oplus}V$ 中相等. 因此,

$$\tau(\sum r_i(u_i,v_i)+S)=\tau(\sum s_i(x_i,y_i)+S),$$

而且τ是线性和满射的. 为了说明τ是单射,我们必须证明如果在$U\overline{\oplus}V$中,$\sum r_i(u_i,v_i)=0$,那么$\sum r_i(u_i,v_i)\in S$.

这样,注意到,作为形式和,$\sum r_i(u_i,v_i)\in S$ 当且仅当运用法则

$$r(u,v)\to(ru,rv),(ru,rv)\to r(u,v),$$

或

$$(u,v)+(u',v')\to(u+u',v+v'),(u+u',v+v')\to(u,v)+(u',v')$$

来替换任何项所得的和也在 S 中. 因此,由$\sum r_i(u_i,v_i)=0$ 通过进行这样的替换,我们

可以把 $\sum r_i(u_i, v_i)$ 简化为 0，而 0 又在 S 中，所以和 $\sum r_i(u_i, v_i)$ 必在 S 中.

因此，τ 是 $\mathcal{F}_{U\times V}/S$ 到 $U \overline{\oplus} V$ 的一个同构.

由此，把 $U \overline{\oplus} V$ 描述为商空间有点勉强. 然而，我们还有它关于可交换图的另一特征. 与直和 $U \overline{\oplus} V$ 相伴的是两个射影 $\rho_1: U \overline{\oplus} V \to U$ 和 $\rho_2: U \overline{\oplus} V \to V$，分别由 $\rho_1((u, v)) = u$ 和 $\rho_2((u, v)) = v$ 定义.

下面我们考虑一下三元组 $(U \overline{\oplus} V, \rho_1, \rho_2)$. 在图 13－4 中，如果 W 是 F 上线性映射为 $f_1: W \to U$ 和 $f_2: W \to V$ 的任意向量空间，那么存在唯一的线性变换 $\tau: W \to U \overline{\oplus} V$，使得图 13－4 可交换，即，使得 $\rho_1 \tau = f_1, \rho_2 \tau = f_2$.

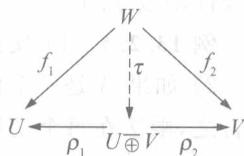


图 13－4

如果这样的一个 τ 存在，那么 $\rho_1(\tau(w)) = f_1(w)$，$\rho_2(\tau(w)) = f_2(w)$. 所以必有

$$\tau(w) = (f_1(w), f_2(w)).$$

这实际上定义了 W 到 $U \overline{\oplus} V$ 的唯一的线性变换. 以下定理说明这个性质表现了直和的特征. 它的证明与定理 13.1.1 的证明很类似.

定理 13.1.2(外直和的泛性质)　令 U, V 是 F 上的向量空间，D 是 F 上的一个向量空间，且令 $\sigma_1: D \to U$ 和 $\sigma_2: D \to V$ 是线性变换，见图 13－5 所示. 设三元组 (D, σ_1, σ_2) 有以下性质：如果 W 是 F 上的任意向量，且如果 $f_1: W \to U$ 和 $f_2: W \to V$ 是线性变换，那么存在唯一的线性变换 $\tau: W \to D$，使得这一图可交换，即，使得 $\sigma_1 \tau = f_1, \sigma_2 \tau = f_2$，那么 D 与外直和 $U \overline{\oplus} V$ 是同构的.

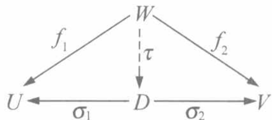


图 13－5

总之，外直和 $U \overline{\oplus} V$ 有以下三个等价的性质：

1) $U \overline{\oplus} V = \{(u, v) \mid u \in U, v \in V\}$.

2) 商空间 $\mathcal{F}_{U\times V}/S$. 这里 $\mathcal{F}_{U\times V}$ 是 $U \times V$ 上的自由向量空间，且

$$S = \operatorname{span}\{\, r(u, v) - (ru, rv), (u, v) + (u', v') - (u + u', v + v') \,\}.$$

3) 定理 13.1.2 给出的外直和的泛性质.

13.2　向量空间的张量积

13.2.1　双线性映射与张量积

在定义张量积之前，我们还需要一些准备.

定义 13.2.1　令 U,V 和 W 是 F 上的向量空间，函数 $f\colon U\times V\to W$ 是**双线性的**，如果它在两个变量中分别是线性的，即，如果

$$f(ru+su',v)=rf(u,v)+sf(u',v);$$
$$f(u,rv+sv')=rf(u,v)+sf(u,v').$$

从 $U\times V$ 到 W 的全体双线性函数所成的集合用 $\mathcal{B}(U,V;W)$ 表示. 基域 F 中有值的双线性函数 $f\colon U\times V\to W$ 叫做 $U\times V$ 上的**双线性型**.

例 13.2.1　1) 实内积 $<\ ,\ >\colon V\times V\to\mathbf{R}$ 是 $V\times V$ 上的一个双线性型.

2) 如果 A 是一个代数，由 $\mu(a,b)=ab$ 定义的积映射 $\mu\colon A\times A\to A$ 是双线性的. 简言之，乘法在每个变量中都是线性的.

如果 V 是一向量空间，从 $V\times V$ 到 W 有两类函数：线性映射 $L(V\times V,W)$ 和双线性映射 $\mathcal{B}(V,V;W)$. 只有零映射同时是这两类映射，即，既是线性又是双线性的映射只有零映射.

现在，我们可以定义两个向量空间的张量积了.

定义 13.2.2　令 U,V 是 F 上的向量空间，且令 T 是自由向量空间 $\mathcal{F}_{U\times V}$ 的子空间，对于 $\forall r,s\in F,u,u'\in U$ 和 $v,v'\in V$，T 由形式为

$$r(u,v)+s(u',v)-(ru+su',v)$$
$$r(u,v)+s(u,v')-(u,rv+sv')$$

的全体向量所生成. 商空间 $\mathcal{F}_{U\times V}/T$ 称为 U 和 V 的**张量积**，记为 $U\otimes V$.

在张量积的这一定义下，如果在每个坐标中，向量空间运算都是线性的，那么我们就除以 $U\times V$ 中等于零的全体向量所生成的空间.

根据这个定义，$U\otimes V$ 中的元素有形式

$$\sum r_i(u_i,v_i)+T.$$

通常，用 $u\otimes v$ 来表示陪集 $(u,v)+T$，因此，$U\otimes V$ 中任意元都有形式 $\sum u_i\otimes v_i$. 这里

$$r(u\otimes v)+s(u'\otimes v)=(ru+su')\otimes v;$$
$$r(u\otimes v)+s(u\otimes v')=u\otimes(rv+sv').$$

因此，

$$\sum u_i\otimes v_i=\sum x_i\otimes y_i,$$

当且仅当应用上面两式进行有限次替换后，我们可以从一种表达式得到另一种表达式.

13.2.2　张量积的泛性质

与外直和相比,虽然这个定义看上去是可行的,但还是有些难以操作,所以我们转而通过一个泛性质来讨论其特征.

定理 13.2.1(张量积的泛性质)　令 U,V 是同一域 F 上的向量空间. $(U\otimes V,t)$ 有以下性质,这里 $t\colon U\times V\to U\otimes V$ 是由 $t(u,v)=u\otimes v$ 定义的双线性映射. 在图 13-6 中,如果 $f\colon U\times V\to W$ 是 F 上 $U\times V$ 到向量空间 W 的任意双线性函数,那么存在唯一的线性变换 $\tau\colon U\otimes V\to W$,使得图 13-6 中的图可交换,即,使得 $\tau\circ t=f$. 而且, $U\otimes V$ 是唯一的,如果 (X,s) 也有这一性质,那么 X 同构于 $U\otimes V$.

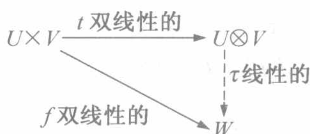


图 13-6

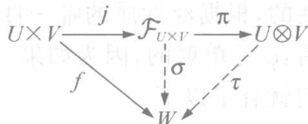


图 13-7

证明　为了证 $(U\otimes V,t)$ 有以上性质,考虑图 13-7 中的图.

因 $t(u,v)=u\otimes v=(u,v)+T$,则映射 $t\colon U\times V\to U\otimes V$ 正好是典范单射 $j\colon U\times V\to\mathcal{F}_{U\times V}$ 的合成,这从典范射影

$$\pi\colon\mathcal{F}_{U\times V}\to U\otimes V=\mathcal{F}_{U\times V}/T$$

中可以得出,即 $t=\pi\circ j$.

由自由向量空间的泛性质可知,存在唯一的 $\sigma\colon\mathcal{F}_{U\times V}\to W$,使得 $\sigma\circ j=f$.

由于 f 是双线性的,它把向量

$$r(u,v)+s(u',v)=(ru+su',v)\text{ 和 }r(u,v)+s(u,v')=(u,rv+sv')$$

中生成 T 的任意向量都映成零向量,所以 $T\subset\operatorname{Ker}(\sigma)$. 因此,我们可知,存在唯一的线性变换 $\tau\colon U\otimes V\to W$,使得 $\tau\circ\pi=\sigma$. 因此,

$$\tau\circ t=\tau\circ\pi\circ j=\sigma\circ j=f.$$

而且,如果 $\tau'\circ t=f$,那么 $\sigma'=\tau'\circ\pi\colon\mathcal{F}_{U\times V}\to W$ 是一个使得

$$\sigma'\circ j(u,v)=\tau'\circ\pi\circ j(u,v)=\tau'\circ t(u,v)=f(u,v)=\sigma\circ j(u,v)$$

的线性变换,所以由 $\sigma'\circ j=\sigma\circ j$,可得 $\sigma'=\sigma$,又可得 $\tau'=\tau$. 因此, τ 是唯一的.

定理 13.2.1 表明,对于每个双线性函数 $f\colon U\times V\to W$,都对应有唯一的线性函数 $\tau\colon U\otimes V\to W$,通过它, f 可以被分解. 这就确定了由 $\varphi(f)=\tau$ 给定的映射 $\varphi\colon\mathcal{B}(U,V;W)\to L(U\otimes V,W)$. 换句话说, $\varphi(f)$ 是唯一的使得

$$\varphi(f): U \otimes V \to W,$$
$$\varphi(f)(u \otimes v) = f(u,v)$$

的线性映射. 注意到 φ 是线性的,因为如果 $f,g \in \mathcal{B}(U,V;W)$,那么

$$[r\varphi(f) + s\varphi(g)](u \otimes v) = rf(u,v) + sg(u,v) = (rf + sg)(u,v),$$

所以由泛性质的唯一性可知

$$r\varphi(f) + s\varphi(g) = \varphi(rf + sg).$$

同样,φ 是满射的,因为如果 $\tau: U \otimes V \to W$ 是任意线性映射,那么

$$f = \tau \circ t: U \times V \to W$$

是双线性的,根据泛性质的唯一性部分,我们有 $\varphi(f) = \tau$.

最后,φ 是单射的,因为如果 $\varphi(f) = 0$,那么 $f = \varphi(f) \circ t = 0$.

我们就有了以下

定理 13.2.2 令 U,V 和 W 是 F 上的向量空间,那么映射

$$\varphi: \mathcal{B}(U,V;W) \to L(U \otimes V, W)$$

是一个同构,且 φ 定义如下:$\varphi(f)$ 是唯一的使得 $f = \varphi(f) \circ t$ 的线性映射. 因此,

$$\mathcal{B}(U,V,W) \cong L(U \otimes V, W).$$

13.2.3 张量积的基本性质

有了张量积的定义和泛性质,我们现在就可以讨论其基本性质了.

定理 13.2.3 如果 $\{u_1,\cdots,u_n\}$ 是 U 中的线性无关向量,$\{v_1,\cdots,v_n\}$ 是 V 中的任意向量,那么由 $\sum u_i \otimes v_i = 0$ 可得对于 $\forall i, v_i = 0$.

证明 考虑到向量 u_i 的对偶向量 $\delta_i \in U^*$. 因此,$\delta_i(u_j) = \delta_{i,j}$. 对于任意线性泛函 $\varepsilon_i: V \to F$,由

$$f(u,v) \to \sum_{j=1}^{n} \delta_i(u)\varepsilon_j(v)$$

定义一双线性型 $f: U \times V \to F$. 那么,根据张量积的泛性质,存在唯一的线性泛函 $\tau: U \otimes V \to F$,使得 $\tau \circ t = f$. 因此,

$$0 = \tau(\sum_i u_i \otimes v_i) = \sum_i \tau \circ t(u_i, v_i) = \sum_i f(u_i, v_i)$$
$$= \sum_i \sum_j \delta_j(u_i)\varepsilon_j(v_i) = \sum_i \varepsilon_i(v_i).$$

由于 ε_i 是任意的,从而推出对于 $\forall i, v_i = 0$.

推论 13.2.4　如果 $u \neq 0, v \neq 0$，那么 $u \otimes v \neq 0$.

定理 13.2.5　令$\mathcal{B} = \{e_i \mid i \in I\}$是$U$的一个基，$C = \{f_j \mid j \in J\}$是$V$的一个基，那么集合$\mathcal{D} = \{e_i \otimes f_j \mid i \in I, j \in J\}$是$U \otimes V$的一个基.

证明　为了说明$\mathcal{D}$是线性无关的，设

$$\sum_{i,j} r_{i,j}(e_i \otimes f_j) = 0，也可以记为\sum_i e_i \otimes (\sum_j r_{i,j} f_j) = 0.$$

所以，对于 $\forall i$，必有$\sum_j r_{i,j} f_j = 0$. 因此，对于 $\forall i, j, r_{i,j} = 0$. 为了说明$\mathcal{D}$生成$U \otimes V$，设$u \otimes v \in U \otimes V$. 由于

$$u = \sum_i r_i e_i, v = \sum_j s_j f_j,$$

所以

$$\begin{aligned} u \otimes v &= \sum_i r_i e_i \otimes \sum_j s_j f_j = \sum_j s_j (\sum_i r_i e_i \otimes f_j) \\ &= \sum_j s_j (\sum_i r_i (e_i \otimes f_j)) = \sum_{i,j} r_i s_j (e_i \otimes f_j). \end{aligned}$$

由于$U \otimes V$中任意向量都是向量$u \otimes v$的一个有限和，所以$\mathcal{D}$生成了$U \otimes V$.

推论 13.2.6　对于有限维向量空间，$\dim(U \otimes V) = \dim(U) \cdot \dim(V)$.

定理 13.2.7　设U, V是有限维向量空间，那么根据$\tau(\alpha \otimes \beta)(u \otimes v) = \alpha(u)\beta(v)$定义的同构

$$\tau: U^* \otimes V^* \to (U \otimes V)^*,$$

$$U^* \otimes V^* \cong (U \otimes V)^*.$$

证明　首先选取$\alpha \in U^*$，$\beta \in V^*$，并考虑由$\sigma_{\alpha,\beta}(u, v) = \alpha(u)\beta(v)$定义的映射$\sigma_{\alpha,\beta}$: $U \times V \to F$. 这一映射是双线性的，所以由张量积的泛性质可知，存在唯一的线性映射$\hat{\sigma}_{\alpha,\beta}$: $U \otimes V \to F$，使得$\hat{\sigma}_{\alpha,\beta}(u \otimes v) = \sigma_{\alpha,\beta}(u, v) = \alpha(u)\beta(v)$. 因而，$\hat{\sigma}_{\alpha,\beta} \in (U \otimes V)^*$. 现在，由$\sigma(\alpha, \beta) = \hat{\sigma}_{\alpha,\beta}$，我们定义映射

$$\sigma: U^* \times V^* \to (U \otimes V)^*.$$

这个映射也是双线性的. 例如，

$$\begin{aligned} \sigma(r\alpha + s\beta, \gamma)(u \otimes v) &= (r\alpha + s\beta)(u)\gamma(v) \\ &= r\alpha(u)\gamma(v) + s\beta(u)\gamma(v) \\ &= r\sigma(\alpha, \gamma)(u, v) + s\sigma(\beta, \gamma)(u, v) \\ &= [r\sigma(\alpha, \gamma) + s\sigma(\beta, \gamma)](u, v), \end{aligned}$$

所以 $$\sigma(r\alpha+s\beta,\gamma)=r\sigma(\alpha,\gamma)+s\sigma(\beta,\gamma).$$

这说明 σ 在它的第一坐标中是线性的. 因此,由泛性质可知存在唯一的线性映射 τ: $U^*\otimes V^*\to(U\otimes V)^*$,使得 $\tau(\alpha\otimes\beta)=\sigma(\alpha,\beta)$,即

$$\tau(\alpha\otimes\beta)(u,v)=\sigma(\alpha,\beta)(u,v)=\hat{\sigma}_{\alpha,\beta}(u\otimes v)=\alpha(u)\beta(v).$$

为了说明 τ 是一个同构,设 $B=\{b_i\}$ 是 U 的一个基,其对偶基为 $B'=\{\beta_i\}$,$X=\{c_i\}$ 是 V 的一个基,其对偶基为 $X'=\{\gamma_i\}$,那么

$$\tau(\beta_i\otimes\gamma_j)(b_u\otimes c_v)=\beta_i(b_u)\gamma_j(c_v)=\delta_{i,u}\delta_{j,v}=\delta_{(i,j),(u,v)}.$$

所以 $\tau(\beta_i\otimes\gamma_j)\in(U\otimes V)^*$ 是属于 $U\otimes V$ 的一个关于基 $\{b_u\otimes c_v\}$ 的对偶基向量. 因此,τ 属于 $U^*\otimes V^*$,并且把基 $\{\tau(\beta_i\otimes\gamma_j)\}$ 变成基 $\{\beta_i\otimes\gamma_j\}$,从而,$\tau$ 是一个同构.

结合定理 13.2.2 和定理 13.2.7 的两个同构,对于有限维向量空间 U 和 V,

$$U^*\otimes V^*\cong(U\otimes V)^*\cong\mathcal{B}(U,V;F).$$

13.3 线性变换的张量积

13.3.1 线性变换的张量积

定义 13.3.1 令 τ: $V\to V'$,σ: $W\to W'$ 是线性变换,那么存在满足

$$(\tau\odot\sigma)(v\otimes w)=\tau(v)\otimes\sigma(w)$$

的唯一的线性变换 $(\tau\odot\sigma)$: $V\otimes W\to V'\otimes W'$. 为了说明这一点,注意到由

$$f(v,w)=\tau(v)\otimes\sigma(w).$$

定义的函数 f: $V\times W\to V'\otimes W'$ 是双线性的,所以根据张量积的泛性质,存在唯一的线性变换 $\tau\odot\sigma$,使得 $(\tau\odot\sigma)(v\otimes w)=\tau(v)\otimes\sigma(w)$ 成立. 映射 $\tau\odot\sigma$ 叫做 τ 和 σ 的**张量积**.

因此,存在一个由 $\varphi(\tau,\sigma)=\tau\odot\sigma$ 定义的映射

$$\varphi\colon L(V,W)\times L(V',W')\to L(V\otimes W)\times L(V'\otimes W').$$

这一映射是双线性的,所以存在满足 $\theta(\tau\otimes\sigma)=\tau\odot\sigma$ 的唯一的线性变换

$$\theta\colon L(V,W)\otimes L(V',W')\to L(V\otimes W,V'\otimes W').$$

我们想要说明 θ 是单射的. 注意到任意非零向量 $\xi\in L(V,W)\otimes L(V',W')$ 都有形式

$$\xi=\sum_{i=1}^{n}\tau_i\otimes\sigma_i,$$

这里 τ_i 和 σ_i 是线性无关的. 为了说明 $\mathrm{Ker}(\theta)=\{0\}$,设

$$\theta(\xi)=\theta(\sum_{i=1}^{n}\tau_i\otimes\sigma_i)=0,$$

那么对于 $\forall v\in V, w\in W, \sum_{i=1}^{n}\tau_i(v)\otimes\sigma_i(w)=0$.

我们选取 $v\in V$,使得 $\tau_1(v)\neq 0$,同时令(有必要的话,通过重新编号)$\tau_1(v),\cdots,\tau_k(v)$ 是 $\tau_1(v),\cdots,\tau_n(v)$ 中的一极大线性无关集. 因此,对于 $u=k+1,\cdots,n$,

$$\tau_u(v)=\sum_{j=1}^{k}r_{u,j}\tau_j(v).$$

从而,有

$$\sum_{i=1}^{n}\tau_i(v)\otimes\sigma_i(w)=0$$

$$\begin{aligned}0&=\sum_{i=1}^{k}\tau_i(v)\otimes\sigma_i(w)+\sum_{u=k+1}^{n}(\sum_{j=1}^{k}r_{u,j}\tau_j(v))\otimes\sigma_u(w)\\&=\sum_{i=1}^{k}\tau_i(v)\otimes\sigma_i(w)+\sum_{j=1}^{k}\tau_j(v)\otimes(\sum_{u=k+1}^{n}r_{u,j}\sigma_u(w))\\&=\sum_{i=1}^{k}\tau_i(v)\otimes(\sigma_i(w)+\sum_{u=k+1}^{n}r_{u,j}\sigma_u(w)).\end{aligned}$$

由于 $\tau_1(v),\cdots,\tau_k(v)$ 是线性无关的,所以对于 $\forall i=1,\cdots,k$ 和 $\forall w\in W$,必有

$$\sigma_i(w)+\sum_{u=k+1}^{n}r_{u,j}\sigma_u(w)=0.$$

因此,

$$\sigma_i+\sum_{u=k+1}^{n}r_{u,j}\sigma_u=0.$$

与 σ_i 是线性无关的这一事实矛盾. 因此,$\theta(\xi)\neq 0$,从而 θ 是单射的.

如果所有向量空间都是有限维的,那么 θ 也是满射的,因此 θ 是一个同构. 由于 $\theta: \tau\otimes\sigma\mapsto\tau\odot\sigma$ 是单射的,所以我们用符号 $\tau\otimes\sigma$ 来代替张量积 $\tau\odot\sigma$.

定理 13.3.1　令 $\tau\in L(V,V')$,$\sigma\in L(W,W')$. 存在唯一的线性变换

$$\tau\odot\sigma\in L(V\otimes W,V'\otimes W'),$$

称为 τ 和 σ 的张量积,它满足 $(\tau\odot\sigma)(v\otimes w)=\tau(v)\otimes\sigma(w)$. 而且,存在满足 $\theta(\tau\otimes\sigma)=\tau\odot\sigma$ 的(唯一的)单射线性变换

$$\theta: L(V,W)\otimes L(V',W')\to L(V\otimes W,V'\otimes W').$$

如果所有向量空间都是有限维的,那么θ是一个同构.

13.3.2　基域的变换

前面我们已经看到,定义在n维实向量空间V上的线性算子τ可能没有n个本征值(数重数),因为它的特征多项式在$\mathbf{R}$上可能不分裂.

另一方面,n维复向量空间上的线性算子却可能有n个本征值.这样,我们就想知道,是否可以把实向量空间扩充为复向量空间,相应地也把实算子扩充为复算子.

为了方便,我们称域F上的向量空间为**F空间**.有几种方法来"升级"向量空间的基域,比如,设V是一个F空间,F'是F的扩张域,即$F' \supset F$.如果$\{b_i\}$是V的一个基,那么V的每一元素x都有形式$x = \sum r_i b_i$,这里$r_i \in F$.

我们只要取形式为$x = \sum r_i' b_i$的全体形式线性组合,就可以定义一个F'空间V',这里$r_i' \in F'$.换句话说,V'是集合$\{b_i\}$上的自由F'空间.

F'空间V'的维数与F空间V的维数是一样的.V'也是一个F空间(只要把数限制在F上),同样把$x \in V$映成$j(x) = x \in V'$的包含映射$j: V \to V'$是一个F单一同态.

上面所描述的方法利用了V中的一个任意选择的基,因此并不是无坐标的.然而,如下所述,我们可以运用张量积给出一种无坐标的方法.如果V是一个F空间,令$V' = F' \otimes_F V$.在表示所取张量积是关于基域F时,通常要包括$\otimes_F$的下标F(所有相关映射都是F双线性或F线性的).然而,由于我们不会取F以外的任何其他域的张量积,所以也就不会一直都用这个符号.

根据张量积的定义,向量空间V'是一个F空间,但如下所述,我们可以使它为一个F'空间.选取$s' \in F'$,考虑由$f_{s'}(r', v) = s'r' \otimes v$定义的映射

$$f_{s'}: (F' \times V) \to (F' \otimes_F V).$$

由于$f_{s'}$是双线性的,所以张量积的泛性质蕴含着存在唯一的F线性映射

$$\tau_{s'}: (F' \otimes_F V) \to (F' \otimes_F V), \text{使得 } \tau_{s'}(r' \otimes v) = s'r' \otimes v.$$

我们可以认为这一映射已乘以了数$s' \in F'$.

注意,由于$\tau_{s'}$是F线性的,所以它是可加的,从而

$$\tau_{s'}(r' \otimes v + u' \otimes w) = \tau_{s'}(r' \otimes v) + \tau_{s'}(u' \otimes w).$$

即
$$s'(r' \otimes v + u' \otimes w) = s'(r' \otimes v) + s'(u' \otimes w).$$

又因为数乘定义的所有性质都是满足的,所以V'确是一个F'空间.不难看出,如果$\{b_i\}$是F空间V的一个基,那么$\{1 \otimes b_i\}$是F'空间V'的一个基,所以F'空间V'的维数与F空间V的维数相同.

由$\upsilon(v) = 1 \otimes v$定义的映射$\upsilon: V \to V'$是一个$F$单一同态,所以$F$空间$V'$中也含

有与 V 相同的一个同构，我们也称 F 线性单一同态 υ 为 V 的 **F' 扩张映射**.

正如以下定理所叙述的一样，这一映射有它自己的一个泛性质.

定理 13.3.2　令 $\upsilon: V \to V' = F' \otimes_F V$ 是 F 空间 V 的 F' 扩张映射，那么 υ 有以下泛性质：对于任意 F 线性映射 $f: V \to W'$，这里 W' 是任意 F' 空间，都存在唯一的 F' 线性映射 $\tau: V' \to W'$，使得图 13-8 中的图是可交换的，即使得 $\tau \circ \upsilon = f$.

证明　映射 $\tau: F' \otimes_F V \to W'$ 存在，那么它必满足

$$\tau(r' \otimes v) = r'\tau(1 \otimes v) = r'f(v).$$

这表明，如果 τ 存在，那么它是由 f 唯一确定的. 考虑由 $g(r', v) = r'f(v)$ 定义的映射 $g: (F' \times V) \to W'$，由于它是双线性的，所以存在使

$$\tau(r' \otimes v) = r'\tau(1 \otimes v) = r'f(v)$$

成立的唯一的 F 线性映射 τ. 容易看出 τ 也是 F' 线性的，因为

$$\tau[s'(r' \otimes v)] = \tau(s'r' \otimes v) = s'r'f(v) = s'\tau(r' \otimes v).$$

对于怎样把 F 线性映射扩充为 F' 线性映射来说，定理 13.3.2 起了重要作用.

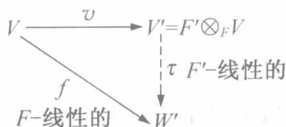


图 13-8



图 13-9

定理 13.3.3　令 V, W 是 F 空间，υ, μ 分别是其 F' 扩张映射，那么对于任意 F 线性映射 $\tau: V \to W$，映射 $\tau' = \iota_{F'} \otimes \tau: V' \to W'$ 使得图 13-9 中图可交换，使得 $\mu \circ \tau = \tau' \circ \upsilon$ 的唯一的 F' 线性映射.

证明　映射 $\mu \circ \tau$ 是 F 空间 V 到 F' 空间 W' 的一个 F 线性映射. 因此，定理 13.3.2 证明了存在唯一的 F' 线性映射 $\tau': V' \to W'$，使得 $\mu \circ \tau = \tau' \circ \upsilon$.

为了说明 $\tau' = \iota_{F'} \otimes \tau$，注意到

$$\begin{aligned}\tau'(r' \otimes v) &= r'\tau(1 \otimes v) = r'(\tau' \circ \upsilon)(v) = r'(\mu \circ \tau)(v) \\ &= r'(1 \otimes \tau(v)) = \iota_{F'}(r') \otimes \tau(v) \\ &= (\iota_{F'} \otimes \tau)(r' \otimes v).\end{aligned}$$

13.3.3　多重线性映射与迭代张量积

张量积运算可以简单地扩充到两个以上的向量空间上.

定义 13.3.2　如果 $V_1, \cdots, V_n$ 和 W 是 F 上的向量空间，函数 $f: V_1 \times \cdots \times V_n \to W$

是多重线性的，如果它在每个变量中都是线性的，即对于 $\forall k=1,\cdots,n$，

$$f(u_1,\cdots,u_{k-1},rv+sv',u_{k+1},\cdots,u_n)=$$
$$rf(u_1,\cdots,u_{k-1},v,u_{k+1},\cdots,u_n)+sf(u_1,\cdots,u_{k-1},v',u_{k+1},\cdots,u_n).$$

n 个变量的多重线性函数也称为 **n 线性函数**.

全体多重线性函数所成的集合记为 $\mathrm{Mul}(V_1,\cdots,V_n;W)$. $V_1\times\cdots\times V_n$ 到基域 F 的多重线性函数称为**多重线性型**(或 **n 型**).

例 13.3.1　行列式函数 $\det: M_n\to F$ 是 M_n 中矩阵的列上的一个 n 线性型.

定义 13.3.3　令 $V_1,\cdots,V_n$ 是 F 上的向量空间，且令对于 $\forall r,s\in F, u,u'\in U$，$v_1,\cdots,v_n\in V$，$T$ 是 $V_1\times\cdots\times V_n$ 上型，形如：

$$r(v_1,\cdots,v_{k-1},u,v_{k+1},\cdots,v_n)+s(v_1,\cdots,v_{k-1},u',v_{k+1},\cdots,v_n)$$
$$-(v_1,\cdots,v_{k-1},ru+su',v_{k+1},\cdots,v_n)$$

的全体向量所生成的自由向量空间 Φ 的子空间. 商空间 Φ/T 称为 $V_1,\cdots,V_n$ 的**张量积**，记作 $V_1\otimes\cdots\otimes V_n$.

与前面一样，我们用 $v_1\otimes\cdots\otimes v_n$ 表示陪集 $(v_1,\cdots,v_n)+T$，所以 $V_1\otimes\cdots\otimes V_n$ 的任意元素都有线性型 $\sum v_{i_1}\otimes\cdots\otimes v_{i_n}$.

张量积也可以由泛性质来描述其特征.

定理 13.3.4(张量积的泛性质)　令 $V_1,\cdots,V_n$ 是域 F 上的向量空间. $(V_1\otimes\cdots\otimes V_n)$ 有以下性质，这里 $t: V_1\times\cdots\times V_n\to V_1\otimes\cdots\otimes V_n$ 是由 $t(v_1,\cdots,v_n)=(v_1\otimes\cdots\otimes v_n)$ 定义的双线性映射. 在图 13-10 中，如果 $f: V_1,\cdots,V_n\to W$ 是 F 上 $V_1,\cdots,V_n$ 到向量空间 W 的任意多重线性函数，那么存在唯一的线性变换

$$\tau: V_1\otimes\cdots\otimes V_n\to W,$$

使得图 13-10 中的图可交换，即使得 $\tau\circ t=f$. 而且，$V_1\otimes\cdots\otimes V_n$ 是唯一的，因为如果 (X,s) 也有这一性质，那么 X 同构于 $V_1\otimes\cdots\otimes V_n$.

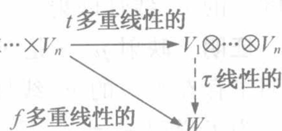


图 13-10

下面是多重张量积的基本性质.

定理 13.3.5　张量积有以下性质(所有向量空间都在域 F 上)：

1)(结合律)　存在同构

$$\tau: (V_1\otimes\cdots\otimes V_n)\otimes(W_1\otimes\cdots\otimes W_m)\to V_1\otimes\cdots\otimes V_n\otimes W_1\otimes\cdots\otimes W_m,$$

使得

$$\tau[(v_1\otimes\cdots\otimes v_n)\otimes(w_1\otimes\cdots\otimes w_m)]=v_1\otimes\cdots\otimes v_n\otimes w_1\otimes\cdots\otimes w_m.$$

特别地，

$$(U \otimes V) \otimes W \cong U \otimes (V \otimes W) \approx U \otimes V \otimes W.$$

2)（交换律）　令 π 是指标$\{1,\cdots,n\}$的任意置换，那么存在同构

$$\sigma: V_1 \otimes \cdots \otimes V_n \to V_{\pi(1)} \otimes \cdots \otimes V_{\pi(n)},$$

使得

$$\sigma(v_1 \otimes \cdots \otimes v_n) = v_{\pi(1)} \otimes \cdots \otimes v_{\pi(n)}.$$

3) 存在同构 $\rho_1: F \otimes V \to V$，使得 $\rho_1(r \otimes v) = rv$；

同理，存在同构 $\rho_2: V \otimes F \to V$，使得 $\rho_2(v \otimes r) = rv$.

因此，$F \otimes V \cong V \cong V \otimes F$.

定理 13.3.5 的类似定理如下：

定理 13.3.6　令 $V_1 \otimes \cdots \otimes V_n$ 和 W 是 F 上的向量空间，那么映射

$$\varphi: \mathrm{Mul}(V_1,\cdots,V_n; W) \to L(V_1 \otimes \cdots \otimes V_n, W)$$

是一个同构，且 φ 是由 $\varphi(f)$ 使得 $f = \varphi(f) \circ t$ 的唯一的线性映射这一事实所定义的，因此，

$$\mathrm{Mul}(V_1,\cdots,V_n; W) \cong L(V_1 \otimes \cdots \otimes V_n, W).$$

而且，如果全体向量空间都是有限维的，那么

$$\dim[\mathrm{Mul}(V_1,\cdots,V_n; W)] = \dim(V_1) \cdots \dim(V_n)\dim(W).$$

13.4　交错映射与外积

13.4.1　交错映射

我们将用符号 V^n 来表示 V 的笛卡儿积，$\otimes^n V$ 表示 n 重张量积.

以下定义描述了多重线性映射的一些特殊类型：

定义 13.4.1　1) 多重线性映射 $f: V^n \to W$ 是**对称**的，如果对于 $\forall i \neq j$，

$$f(v_1,\cdots,v_i,\cdots,v_j,\cdots,v_n) = f(v_1,\cdots,v_j,\cdots,v_i,\cdots,v_n).$$

2) 多重线性映射 $f: V^n \to W$ 是**斜对称**的，如果对于 $i \neq j$，

$$f(v_1,\cdots,v_i,\cdots,v_j,\cdots,v_n) = -f(v_1,\cdots,v_j,\cdots,v_i,\cdots,v_n).$$

3) 多重线性映射 $f: V^n \to W$ 是**交错**的，当任意两个向量 v_i 相等时，

$$f(v_1,\cdots,v_n) = 0.$$

现在,我们讨论一下置换问题.集合 $N=\{1,\cdots,n\}$ 的**置换**是双射函数 $\pi: N\rightarrow N$.我们用 S_n 来表示所有这种置换所成的集合.这是 n 个符号上的对称群.长度为 k 的轮换是一个形为 $(i_1,i_2,\cdots,i_k)$ 的置换,它把 i_1 换为 i_2,i_2 换为 i_3,$\cdots$,i_{k-1} 换为 i_k,i_k 换为 i_1(我们假设对于 $u\neq v, i_u\neq i_v$).N 的其他元素都是不动的.

每一个置换都是不相交轮换的积(合成).

转置是长度为2的轮换 (i,j).每个轮换(从而每个置换)都是转置的积.通常,一个置换可以用多种方式表示为转置的积.但是,无论以积的何种形式表示出一给定置换,转置数总是为偶数或奇数.因此,在一个由 π 作为转置的积的任意分解中,我们可以把置换 $\pi\in S_n$ 的**奇偶性**定义为转置数的奇偶性.置换的正负号由 $sg(\pi)$ 定义.因此,如果 π 是一个偶置换,那么 $sg(\pi)=1$;如果 π 是一个奇置换,那么 $sg(\pi)=-1$.π 的正负号常记作 $(-1)^\pi$.

对于所有置换 $\pi\in S_n$,f 是对称的,当且仅当 $f(v_1,\cdots,v_n)=f(v_{\pi(1)},\cdots,v_{\pi(n)})$.

对于所有置换 $\pi\in S_n$,f 是交错的,当且仅当 $f(v_1,\cdots,v_n)=(-1)^\pi f(v_{\pi(1)},\cdots,v_{\pi(n)})$.

如果 f 是一个多重线性函数,那么

$$f(v_1,\cdots,v_i+v_j,\cdots,v_i+v_j,\cdots,v_n)=f(v_1,\cdots,v_i,\cdots,v_i,\cdots,v_n)+$$

$$f(v_1,\cdots,v_i,\cdots,v_j,\cdots,v_n)+f(v_1,\cdots,v_j,\cdots,v_i,\cdots,v_n)+f(v_1,\cdots,v_j,\cdots,v_j,\cdots,v_n).$$

因此,若 f 是交错的,则它也是斜对称的.另一方面,若 f 是斜对称的,

$$f(v_1,\cdots,v_i,\cdots,v_i,\cdots,v_n)=-f(v_1,\cdots,v_i,\cdots,v_i,\cdots,v_n).$$

如果 $\mathrm{char}(F)\neq 2$,那么从中得出 $f(v_1,\cdots,v_i,\cdots,v_i,\cdots,v_n)=0$.

因此,f 是交错的.

13.4.2 外积的泛性质

定义 13.4.2 令 V 是域 F 上 $\mathrm{char}(F)\neq 2$ 的一向量空间,且令 $\bigotimes^n V$ 是 V 自身的 n 重张量积.对于 $\forall i>j$,U 是 $\bigotimes^n V$ 的形式为

$(v_1\otimes\cdots\otimes v_i\cdots\otimes v_j\otimes\cdots\otimes v_n)+(v_1\otimes\cdots\otimes v_j\cdots\otimes v_i\otimes\cdots\otimes v_n)$

的全体元素所成的子空间.商空间 $(\bigotimes^n V)/U$ 称为 V 的**第 n 个外积空间**,记作

$$\wedge^n V \text{ 或 } \underbrace{V\wedge\cdots\wedge V}_{n\text{个因数}}.$$

我们用 $v_1\wedge\cdots\wedge v_n$ 表示陪集 $(v_1\otimes\cdots\otimes v_n)+U$,称 $\wedge$ 为**楔积**.

因此,$v_1\wedge\cdots\wedge v_n$ 的任意元素都有形式 $\sum v_{i_1}\wedge\cdots\wedge v_{i_n}$.

这里每个变量的向量空间运算都是线性的,任意两个变量的互换都会出现一个负号.

外积也可以用泛性质来表现其特征.

定理 13.4.1(外积的泛性质)　令 $V_1,\cdots,V_n$ 是域 F 上 $\mathrm{char}(F)\neq 2$ 的向量空间. $(V_1\wedge\cdots\wedge V_n,a)$ 有以下性质：这里

$$a: V_1\times\cdots\times V_n\rightarrow V_1\wedge\cdots\wedge V_n$$

是由

$$a(v_1,\cdots,v_n)=v_1\wedge\cdots\wedge v_n$$

定义的交错多重线性映射. 在图 13－11 中，如果

$$f: V_1\times\cdots\times V_n\rightarrow W$$

是 F 上 $V_1\times\cdots\times V_n$ 到向量空间 W 的任意交错多重线性函数，那么存在使得图 13－11 中的图可交换，即，使得 $\tau\circ a=f$ 的唯一的线性变换



图 13－11

$$\tau: V_1\wedge\cdots\wedge V_n\rightarrow W,$$

而且，如果(X,σ) 也有这一性质，那么 X 同构于 $V_1\wedge\cdots\wedge V_n$，从而 $V_1\wedge\cdots\wedge V_n$ 是唯一的.

第 14 章　仿射几何与多项式函数

在本章中，我们将研究有限维向量空间 V 的几何，包括仿射几何、平坦格、仿射变换、射影几何以及它保持映射时的结构. 我们还要深入讨论一下多项式函数、形式幂级数、当作线性算子的形式幂级数、连续算子、算子伴随等内容.

14.1　格代数基础

在有关群、环的讨论中，我们已经看到，不仅涉及群、环中的元素的二元合成，而且常常要讨论关于子群、正规子群、理想等的二元合成. 如 I_1, I_2 是环 $\mathbf{R}$ 的两个理想，我们要讨论它们的"交" $I_1 \cap I_2$ 及"和" $I_1 + I_2$ 等. 所有这些都导致了另一类具有两个二元合成的代数结构 —— 格. 格的概念是由 Dedekind(1831—1916) 首先引入的，但当时并没有引起人们多大的重视，直到大约 1930 年，人们才逐渐认识到格理论在代数学的各个分支中的重要性，从而对格作了广泛深入的研究.

14.1.1　格的定义及基本性质

定义 14.1.1　一个偏序集 L 叫做**格**，是指对任何 $a,b \in L$，存在 $\{a,b\}$ 的最小上界 $l.u.b.(X)$ 及最大下界 $g.l.b.(X)$，分别以 $a \vee b$ 及 $a \wedge b$ 记之.

若 L 是一个格，$a,b,c \in L$，考虑 $(a \vee b) \vee c$，它是 $\{a,b,c\}$ 的上界，而且是最小上界，从而 $(a \vee b) \vee c = l.u.b.(\{a,b,c\})$. 同样 $(a \wedge b) \wedge c = g.l.b.(\{a,b,c\})$.

定义 14.1.2　格 L 叫**完全格**，是指对任何非空子集 X，$g.l.b.(X)$ 及 $l.u.b.(X)$ 都存在.

设 L 是一个格，若 L 有最大元，称这个最大元为 L 的**全元素**，常记为 1，于是对任何 $a \subseteq L$，有 $a \leqslant 1$；若 L 有最小元，称这个最小元为 L 的**零元素**，常记为 0，于是对任何 $a \subseteq L$，有 $0 \leqslant a$.

若 L 是完全格，则显然有 $l.u.b.(L) = 1, g.l.b.(L) = 0$.

例 14.1.1　设 S 是任一集，$P(S)$ 是 S 的幂集，偏序集 $(P(S), \subseteq)$ 是一个格，对 $X, Y \in P(S)$，$X \wedge Y = X \cap Y, X \vee Y = X \cup Y$，它还是一个完全格，$1 = S, 0 = \varnothing$.

例 14.1.2　设 G 是一个群，$L(G)$ 表示 G 的所有子群作成的集合，则 $L(G)$ 关于集合

包含关系"$\subseteq$"作成一个偏序集，任取 $H_1, H_2 \in L(G)$，用 $(H_1 \cup H_2)$ 表示 H_1, H_2 生成的子群，则 $(H_1 \cup H_2)$ 是偏序集 $(L(G), \subseteq)$ 中 $\{H_1, H_2\}$ 的最小上界，即 $H_1 \vee H_2 = (H_1 \cup H_2)$，而 $H_1 \cap H_2$ 是 $\{H_1, H_2\}$ 的最大下界，即 $H_1 \wedge H_2 = H_1 \cap H_2$，从而 $L(G)$ 关于偏序 $\subseteq$ 作成一个格，称这个格为 G 的子群格. 它也是一个完全格，$1 = G, 0 = \{e\}$.

例 14.1.3　设 R 是任意环，则 R 的一切子环所成集合以及一切理想所成集合关于集合的包含关系都作成格，两个子环的最大下界是其交，而最小上界是这两个子环生成的子环.

两个理想的最大下界是其交，而最小上界是其和. 考虑环 **R** 的所有理想的集，对 **R** 的理想 I_1, I_2，有 $I_1 \wedge I_2 = I_1 \cap I_2, I_1 \vee I_2 = I_1 + I_2$，它也是一个完全格，$1 = R, 0 = \{0\}$.

例 14.1.4　设 V 是域 F 上的向量空间，S 是 V 的所有子空间作成的集合，则 S 关于集合的包含"$\subseteq$"作成一个偏序集，由于两个子空间的和与交仍是子空间，故 S 中任意两个元都有最小上界与最大下界，从而 $(S, \subseteq)$ 是一个格，称这个格为 V 的子空间格.

对 V 的子空间 V_1, V_2，有 $V_1 \wedge V_2, V_1 \cap V_2, V_1 \vee V_2 = V_1 + V_2$，它也是一个完全格，$1 = V, 0 = \{0\}$.

定理 14.1.1　若偏序集 S 有最大元 1，且 S 的每一子集 $X \neq \varnothing$ 有最大下界，则 S 是一个完全格；对偶地，若偏序集 S 有最小元 0，且对每一子集 $X \neq \varnothing$ 有最小上界，则 S 是一个完全格.

证明　对第一个结论，我们来证明 X 有最小上界. 事实上，令 $Y = \{a \mid a \in S, a$ 是 X 的上界$\}$，则 $1 \in Y$，故 $Y \neq \varnothing$，从而 Y 有最大下界，设 b 是 Y 的最大下界，则对每一 $x \in S$，都是 Y 的下界，从而有 $x \leqslant b$；又若 u 是 X 的上界，则 $u \in Y$，但 b 是 Y 的下界，故必有 $b \leqslant u$，从而 b 是 X 的最小上界. 因此 S 是完全格.

推论 14.1.2　若 L_1, L_2 是两个格，则 $L_1 \times L_2$ 也是格.

14.1.2　格代数

在 L 中，对任意 $a, b \in L$，存在 $a \wedge b = g.l.b.(\{a,b\})$ 及 $a \vee b = l.u.b.(\{a,b\})$. 我们自然可以把 $\wedge$ 及 $\vee$ 看作是集 L 上的两种二元合成.

下面讨论这两种二元合成的一系列性质，从而得到格作为一种代数结构的又一定义.

设 L 是一个格，若 $a, b, c \in L$，则二元合成 $\wedge$，$\vee$ 具有下述性质：

L_1：交换律　$a \wedge b = b \wedge a, a \vee b = b \vee a$；

L_2：结合律　$(a \wedge b) \wedge c = a \wedge (b \wedge c), (a \vee b) \vee c = a \vee (b \vee c)$；

L_3：幂等性　$a \wedge a = a, a \vee a = a$；

L_4：吸收律 $(a \wedge b) \vee a = a, (a \vee b) \wedge a = a$.

定义 14.1.3 设 L, L' 是两个格，映射 $\theta: L \to L'$ 叫做**保序**的，是指对任意 $a, b \in L$，$a \leqslant b \Leftrightarrow a\theta \leqslant b\theta$.

定理 14.1.3 设 L 是一个格，对 $a \in L$，映射 $\theta: L \to L$，对 $x \in L$，$xu = x \wedge a$ 是保序的；映射 $\eta: L \to L$，对 $x \in L$，$x\eta = x \vee a$ 也是保序的.

关于 η 的保序性可以类似地证明

定理 14.1.4 在格 L 中，对任何 $a, b, c \in L$，下述两个不等式成立：

$$(a \wedge b) \vee (a \wedge c) \leqslant a \wedge (b \vee c), \tag{14.1}$$

$$a \vee (b \wedge c) \leqslant (a \vee b) \wedge (a \vee c). \tag{14.2}$$

证明 关于(14.1)式，因为 $b \leqslant b \vee c$，故 $a \wedge b \leqslant a \wedge (b \vee c)$；同样由 $c \leqslant b \vee c$ 得 $a \wedge c \leqslant a \wedge (b \vee c)$，从而得 $(a \wedge b) \vee (a \wedge c) \leqslant a \wedge (b \vee c)$.

关于(14.2)式可类似证明. 这两个不等式通常叫做**分配律不等式**.

推论 14.1.5 在格 L 中，$a \leqslant c \Rightarrow a \vee (b \wedge c) \leqslant (a \vee b) \wedge c$.

事实上，只需在(14.2)式中，当 $a \leqslant c$ 时 $a \vee c = c$ 即可得到. 这个不等式常叫**模不等式**.

定理 14.1.6 设 L 是一个非空集，在 L 中定义两个二元合成 $\wedge$ 及 $\vee$，满足定理中的 $L_1 \sim L_4$. 今在 L 中定义关系"$\leqslant'$"如下：$a \leqslant' b$ 是指 $a \wedge b = a$，则 $\leqslant'$ 是一个偏序关系，偏序集 $\{L, \leqslant'\}$ 是一个格，在这个格中，$a \wedge b$ 正是 $\{a, b\}$ 的最大下界，$a \vee b$ 正是 $\{a, b\}$ 的最小上界.

定义 14.1.4 一个格是具有二元合成 $\wedge$，$\vee$ 的代数结构 $(L, \wedge, \vee)$，它满足定理 4.2.1 中所述的条件 $L_1 \sim L_4$.

对偶原理 若命题 P 是可以从格的公理推出的命题，则它的对偶命题也是可以从格的公理推出的.

由于我们把格看作是具有两个二元合成的一种代数结构，于是与群、环的情形一样，有下述格的同态、同构概念.

定义 14.1.5 格 L 到格 L' 的映射 θ 叫**格的同态**或简称**同态**，是指对任何 $x, y \in L$ 有 $(x \wedge y)\theta = x\theta \wedge y\theta$，$(x \vee y)\theta = x\theta \vee y\theta$.

如果同态 θ 还是双射，则称 θ 为**同构**. 当存在同构 $\theta: L \to L'$ 时，称格 L 同构于 L'，记作 $L \cong L'$. 若 $\theta: L_1 \to L_2$，$\eta: L_2 \to L_3$ 都是格的同态，则 $\theta: L_1 \to L_3$ 也是同态. 若 θ, η 都是同构，则 $\theta\eta$ 也是同构，$\theta^{-1}: L_2 \to L_1$ 也是同构.

由于格中的二元合成 $\wedge$，$\vee$ 与序关系 $\leqslant$ 的内在联系，是否可用 $\leqslant$ 来刻画一个映射是同构呢？

定理 14.1.7 双射 $\theta: L \to L'$ 是格的同构 $\Leftrightarrow \theta, \theta^{-1}$ 都是保序的.

14.1.3　分配格

在一般的格中，对二元合成 $\wedge$ 及 $\vee$ 两者之间除了吸收律这一限制外并没有什么别的要求. 如果对它们之间的联系加以更严格的限制，便得到各种特殊的格. 一种很自然的限制就是所谓分配律，即对格 L 中的任何元素 x,y,z，有

$$L_5 \quad a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c), \tag{14.3}$$

$$a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c), \tag{14.4}$$

如果我们把"$\wedge$"看作乘法，"$\vee$"看作加法，则它们正如通常所说的，是乘法对加法的分配律. 值得注意的是，在一个格中，(14.3) 式与(14.4) 式是等价的，即在任意格 L 中，只要两个分配律之一成立，那么另一分配律一定也成立.

定义 14.1.6　格$(S, \vee, \wedge)$说是一个**分配格**，若 $\forall a,b \in S$，下面算律成立：

$$L_5 \quad a \wedge (b \vee c) = (a \wedge b) \vee (a \wedge c), a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c).$$

例 14.1.5　Z 关于整除所作成的格是分配格，因为，

$$\forall a,b \in \mathbf{Z}, [a,(b,c)] = ([a,b],[a,c])，即\ a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c).$$

例 14.1.6　考虑实数域 $\mathbf{R}$ 上的二维向量空间 V，设 e_1, e_2 是 V 的一组基，考虑 V 的子空间 $Re_1 = \{re_1 \mid r \in \mathbf{R}\}$，$Re_2 = \{re_2 \mid r \in \mathbf{R}\}$，$R(e_1+e_2) = \{r(e_1+e_2)\} \mid r \in \mathbf{R}\}$，则

$$R(e_1+e_2) \wedge (Re_1 \vee Re_2) = R(e_1+e_2) \wedge 1 = R(e_1+e_2),$$

但

$$(R(e_1+e_2) \wedge Re_1) \vee (R(e_1+e_2) \wedge Re_2) = 0 \vee 0 = 0,$$

从而 V 的子空间格不是分配格.

定理 14.1.8　设 L 是分配格，$a,x,y \in L$，若 $a \wedge x = a \wedge y, a \vee x = a \vee y$，则 $x = y$.

14.1.4　模格

我们讨论比分配格更广泛的模格.

定义 14.1.7　一个格 L 叫做**模格**，是指在这个格中下述条件成立：

M：若 $a \leqslant c$，则 $a \vee (b \wedge c) = (a \vee b) \wedge c$. 由于 $a \leqslant c$ 时，$a \vee c = c$，从而分配格是模格.

条件 M 的对偶是 M'：若 $c \leqslant a$，则 $a \wedge (b \vee c) = (a \wedge b) \vee c$.

显然，M' 就是 M，从而在模格中对偶原理仍然有效.

下面的定理指出了模格的重要性：

定理 14.1.9　设 G 是一个群，则 G 的一切正规子群作成的格 $N(G)$ 是一个 Dedekind 格.

由于这个定理,并注意到模格的子格是模格,从而一个环的理想所成的格,左(右)理想所成的格,数域 P 上的向量空间的子空间所成的格等都是模格.

定理 14.1.10　格 L 是模格的充要条件是对 $a,b\in L,a\leqslant b$,且对某一 $c\in L$ 有 $a\vee c=b\vee c,a\wedge c=b\wedge c$,则 $a=b$.

定理 14.1.11　设 L 是 Dedekind 格,$a,b\in L$,则子格 $a\vee b/a$ 同构于 $b/a\wedge b$.

定义 14.1.8　设 $a,b\in L,a<b$,若 L 中不存在 x,使 $a<x<b$,则说 b **覆盖** a.

设 $x\in L$,若在 L 中存在有限序列 $x=x_0,x_1,x_2\cdots x_n=0$,使

$$x=x_0>x_1>\cdots>x_n=0,\tag{14.5}$$

并且 x_i 覆盖 $x_{i+1}(i=0,1\cdots n-1)$,则说(14.5)式是 x 到 0 的一个**极大链**,d 是(14.5)式的长度. x 到 0 的所有极大链的长度中最大者叫做 x 的**维数**,用符号 $d(x)$ 表示.

在域 F 上向量空间 V 的子空间格 $L(V)$ 中,我们知道,若 x 是有限维子空间,则 x 到 0 的每一极大链都有相同长度,即 x 的维数.

定理 14.1.12　设 L 是有零元的 Dedekind 格,$x\in L$,若 x 的维数 $d(x)$ 有限,则 x 到 0 的每一极大链的长度都相等.

同样,对环 R,把上面 G 的正规子群换成 R 的理想,即得到环 R 的主列的概念. 一个群 G(或环 R) 若有主列,则任意两个主列都有相同的长度.

域 F 上向量空间的维数定理:$\dim W_1+\dim W_2=\dim(W_1+W_2)+\dim(W_1\cap W_2)$.

这个事实对一般 Dedekind 格也成立,即

定理 14.1.13(维数定理)　设 L 是有零元的 Dedekind 格,$a,b\in L$,若 $d(a),d(b)$ 均为有限,则 $d(a)+d(b)=d(a\vee b)+d(a\wedge b)$.

证明　在 L 中,子格 $a\vee b/a$ 与 $b/a\wedge b$ 同构,故维数相同. 而 $a\vee b/a$ 的维数等于 $d(a\vee b)-d(a)$,$b/a\wedge b$ 的维数等于 $d(b)-d(a\wedge b)$,故有 $d(a\vee b)-d(a)=d(b)-d(a\wedge b)$,从而 $d(a)+d(b)=d(a\vee b)+d(a\wedge b)$.

维数定理是 Dedekind 格的一个特征,定理的逆命题也成立.

定理 14.1.14　设 L 是有零元的格,且每一个元的维数均有限,对 $\forall a,b\in L$,有 $d(a)+d(b)=d(a\vee b)+d(a\wedge b)$,那么 L 是一个 Dedekind 格.

14.1.5　有余模格

定义 14.1.9　设格 S 有单位元 I 和零元 0,取定 $a\in S$,若存在 $x\in S$,使 $a\wedge x=0,a\vee x=I$,则说 x 是 a 的一个**补元**,记为 a'.

在分配格中,若 a 有补元 a',则只能有一个,因为,假定 a 有两个补元 x,y,由 $a\wedge x=a\wedge y=0,a\vee x=a\vee y=I$,可得 $x=y$.

由补元定义可知,若 a 的补元为 a',则 a' 的补元为 a,即 $(a')'=a$,因此我们也说 a

与 a' 互为补元.

定义 14.1.10 具有零及全元素的格 L 叫做**有余格**,是指任何 $a \in L$ 都有余元素.

例 14.1.7 对任何集 S,格 $\{P(S), \subseteq\}$ 是一个有余格,对 $X \in P(S)$,$X' = \{a \mid a \in S, a \notin X\}$ 是 X 的余元素(空集 $\varnothing$ 是 $P(S)$ 的零元素,S 是 $P(S)$ 的全元素).

在有余格中一个元素的余元素不一定是唯一的. 例如,设 $\{e_1, e_2\}$ 是 V 的一组基,令 $S = Re_1$,则 $T_1 = Re_2$ 及 $T_2 = R(e_1 + e_2)$ 都是 S 的余元素,但 $T_1 \neq T_2$.

定理 14.1.15 在有单位元和零元的分配格中,一切补元的集合作成一个子格.

定理 14.1.16 若 L 是有余模格,$a \in L$,令 $L_a = \{x \mid x \in L, x \leqslant a\}$,则 L_a 也是有余模格.

当 L 是有余格而不是模格时,L_a 不一定是有余格. 例如,在格 N_5 中,$L_x = \{0, Y, X\}$ 不是有余格,其中 Y 没有余元素.

14.2 仿射几何

14.2.1 仿射几何

定义 14.2.1 令 V 是一向量空间,若 $v \in V$,S 是 V 的一子空间,则集合 $v + S = \{v + s \mid s \in S\}$ 叫做 V 的一个**平坦集**,或**陪集**. V 中全体平坦集所成的集合 $\mathcal{A}(V)$ 叫做 V 的**仿射几何**. $\mathcal{A}(V)$ 的**维数** $\dim(\mathcal{A}(V))$ 定义为 $\dim(V)$.

显然,V 的平坦集不过是 V 的平移子空间. 用字母 $S, T, \cdots$ 表示 V 的子空间,$X, Y, \cdots$ 表示 V 的平坦集. 以下是平坦集的一些基本性质:

定理 14.2.1 1) 以下各条等价:

a) $x + S = y + S$; b) $x \in y + S$; c) $x - y \in S$.

令 $X = x + S, Y = y + T$ 是 V 的平坦集,那么

2) $S \subset T \Leftrightarrow$ 对于 $\forall v \in V, v + X \subset Y$.

3) $S = T \Leftrightarrow$ 对于 $\forall v \in V, v + X = Y$.

4) $X \cap Y \neq \varnothing, S \subset T \Rightarrow X \subset Y$.

5) $X \cap Y \neq \varnothing, S = T \Rightarrow X = Y$.

证明 2) 注意到 $S = -x + X, T = -y + Y$,所以

$$S \subset T \Leftrightarrow -x + X \subset -y + Y \Leftrightarrow (y - x) + X \subset Y.$$

3) 因为 $(y - x) + X \subset Y, (x - y) + Y \subset X$,所以 $(y - x) + X \subset Y \subset (y - x) + X$,则 $(y - x) + X = Y$.

4) 设 $z \in X \cap Y$. 由论断 2) 可知,$v + X \subset Y$,所以 $v + z = y \in Y$,则 $v = y - z \in T$. 因此 $X \subset -v + Y \subset Y$.

5）可由 4）直接得出.

由定理,一个平坦集有多种表示法,比如 $x+S$. 当一个平坦集记为 $x+S$ 时,我们称 x 为**平坦代表**,或平坦集的**陪集代表**. 一个平坦集的任意元都可以作为一平坦代表. 另一方面,当 $v=0$ 时,每一平坦集 $x+S$ 都与唯一子空间 S 相伴.

定义 14.2.2 平坦集 $x+S$ 的**维数**是 $\dim(S)$. 维数为 k 的平坦集称为 ***k* 平坦集**.

0 平坦集是一个**点**,1 平坦集是一条**直线**,2 平坦集是一个**平面**. 维数为 $\dim(A(V))-1$ 的平坦集称为**超平面**.

定义 14.2.3 两个平坦集 $X=x+S$ 和 $Y=y+T$ 称为是**平行的**,如果 $S\subset T$,或 $T\subset S$,记为 $X\parallel Y$.

根据定理 14.2.1,若 $X\parallel Y$,则 $X\subset Y$,$Y\subset X$ 或 $X\cap Y=\varnothing$,而且,X 和 Y 是平行的 $\Leftrightarrow$ 这些平坦集中,任意一个平坦集的某一平移包含在另一个平坦集之中.

14.2.2 仿射组合

如果 $r_i\in F$,$r_1+\cdots+r_n=1$,那么线性组合 $r_1x_1+\cdots+r_nx_n$ 称为向量 $x_1,\cdots,x_n$ 的**仿射组合**.

定理 14.2.2 如果 $\operatorname{char}(F)\neq 2$,则对于 V 的子集 X,以下各条等价:

1）在取 X 的任意两点的仿射组合下 X 是闭的,即 $x,y\in X$ 有 $rx+(1-r)y\in X$.

2）在取仿射组合的条件下,X 是闭的,即

$$x_1,\cdots,x_n\in X,r_1+\cdots+r_n=1\Rightarrow r_1x_1+\cdots+r_nx_n\in X.$$

证明 显然,由论断 2）可得 1）. 至于逆命题,我们用归纳法. 由论断 1）,对于 $x_1,x_2\in X$,由于 $r_1+r_2=1$,所以 $r_1x_1+r_2x_2\in X$. 为了便于归纳,设对于 $x_i\in X$,

$$r_1+\cdots+r_{n-1}=1\Rightarrow r_1x_1+\cdots+r_{n-1}x_{n-1}\in X.$$

令 $x_1,\cdots,x_n\in X$,$r_1+\cdots+r_n=1$,考虑仿射组合 $z=r_1x_1+\cdots+r_nx_n$. 如果 r_1 或 r_2 中有一个不为 1,设 $r_1\neq 1$,那么我们可以记为

$$z=r_1x_1+(1-r_1)\left(\frac{r_2}{1-r_1}x_2+\cdots+\frac{r_n}{1-r_1}x_n\right).$$

由于 $\frac{r_2}{1-r_1}x_2+\cdots+\frac{r_n}{1-r_1}x_n$ 的系数之和为 1,所以由归纳假设知 $\frac{r_2}{1-r_1}x_2+\cdots+\frac{r_n}{1-r_1}x_n$ 在 X 中,那么由论断 1）得,$z\in X$. 另一方面,如果 $r_1=r_2=1$,那么由于 $\operatorname{char}(F)\neq 2$,我们可以记

$$z=2\left(\frac{1}{2}x_1+\frac{1}{2}x_2\right)+r_3x_3+\cdots+r_nx_n,$$

而$\frac{1}{2}x_1+\frac{1}{2}x_2\in X$,所以可以再由归纳假设推出,$z\in X$. 无论哪种情形,$z\in X$,所以论断 2) 成立.

$\mathrm{char}(F)\neq 2$ 这个条件是必要的,因为子集 $X=\{(0,0),(1,0),(0,1)\}$ 满足论断 1),但不满足论断 2). 现在我们可以描述平坦集的特征了.

定理 14.2.3　1) 在 V 中,V 的子集 X 是一个平坦集 $\Leftrightarrow$ 在取仿射组合下,X 是闭的 $\Leftrightarrow x_1,\cdots,x_n\in X, r_1+\cdots+r_n=1\Rightarrow r_1x_1+\cdots+r_nx_n\in X$.

2) 如果 $\mathrm{char}(F)\neq 2$,V 的子集 X 是一个平坦集,当且仅当 X 包含穿过其任意两点的直线,即当且仅当 $x,y\in X\Rightarrow rx+(1-r)y\in X$.

证明　设 $X=x+S$ 是一个平坦集,$x_1,\cdots,x_n\in X$,那么对于 $s_i\in S, x_i=x+s_i$,所以如果 $\sum r_i=1$,那么 $\sum r_ix_i=\sum r_i(x+s_i)=x+\sum r_is_i\in x+S$,所以 X 在仿射组合下是闭的. 反之,在仿射组合条件下 X 是闭的,且对于 $\forall x_0\in X$,令 $S=\{x-x_0\mid x\in X\}$. 如果 $x_1-x_0,\cdots,x_n-x_0$ 是 S 中任意向量,$r_1+\cdots+r_n\in F$,那么

$$\sum r_i(x_i-x_0)=r_1x_1+\cdots+r_nx_n+(1-r_1-\cdots-r_n)x_0-x_0\in S.$$

因此,在取线性组合的条件下,S 是闭的,所以 S 是 V 的一个子空间,则 $X=x_0+S$ 是一个平坦集.

2) 可以从 1) 中得到.

14.2.3　仿射包

以下定义给出了向量集所生成的子空间的类似定义.

定义 14.2.4　令 W 是 V 中向量所成的一非空集. W 的**仿射包** $\mathrm{hull}(W)$ 是 W 中最小的平坦集. 我们也称 $\mathrm{hull}(W)$ 为由 W 生成的平坦集.

定理 14.2.4　令 W 是 V 的任意非空子集,仿射包 $\mathrm{hull}(W)$ 是 W 中向量的全体仿射组合而成的集合,$\mathrm{hull}(W)=\{\sum_{i=1}^{n}r_ix_i\mid n\geqslant 1, x_1,\cdots,x_n\in W,\sum_{i=1}^{n}r_i=1\}$.

证明　根据定理 14.2.3,任意包含 W 的平坦集必包含 W 中向量所成的全体仿射组合. 因此,只要说明所有这样的仿射组合所成的集合 X 是一个平坦集,从而,令 $y\in X$,考虑集合 $S=\{y_j-y\mid y_i\in X\}$,则 S 是 V 的一子空间,那么 $X=y+S$ 确是一平坦集.

这样,令 $y=\sum_{i=1}^{n}r_{0,i}x_i, y_1=\sum_{i=1}^{n}r_{1,i}x_i, y_2=\sum_{i=1}^{n}r_{2,i}x_i$.

因此,y_1-y 和 y_2-y 的任意线性组合都有形式

$$z=s(y_1-y)+t(y_2-y)=s\sum_{i=1}^{n}r_{1,i}x_i+t\sum_{i=1}^{n}r_{2,i}x_i-(s+t)y$$

$$= \sum_{i=1}^{n}(sr_{1,i}+tr_{2,i})x_i-(s+t-1)y-y$$

$$= \sum_{i=1}^{n}(sr_{1,i}+tr_{2,i}-(s+t-1)r_{0,i})x_i-y.$$

但是$\sum_{i=1}^{n}(sr_{1,i}+tr_{2,i}-(s+t-1)r_{0,i})x_i = s\sum_{i=1}^{n}r_{1,i}+t\sum_{i=1}^{n}r_{2,i}-(s+t-1)\sum_{i=1}^{n}r_{0,i}$

$$= s+t-(s+t-1)=1.$$

这说明 $z\in S$. 因此,S 是 V 的一个子空间.

向量所成的一个有限集的仿射包记为 $\mathrm{hull}\{x_1,\cdots,x_n\}$. 至于证明

$$\mathrm{hull}\{x_1,\cdots,x_n\}=x_i+\langle x_1-x_i,\cdots,x_{i-1}-x_i,x_{i+1}-x_i,\cdots,x_n-x_i\rangle,$$

这里$\langle x_1-x_i,\cdots,x_{i-1}-x_i,x_{i+1}-x_i,\cdots,x_n-x_i\rangle$是尖括号中向量所生成的一个子空间,则 $\dim(\mathrm{hull}\{x_1,\cdots,x_n\})\leqslant n-1$. 一对互异点所成仿射包是穿过这些点的直线,记作

$$\overline{xy}=\{rx+(1-r)y\mid r\in F\}=y+\langle x-y\rangle.$$

14.3 平 坦 格

14.3.1 平坦格

由于平坦集是 V 的子集,所以它们是以集合的包含关系为偏序的.

定理 14.3.1 V 中平坦集所成的非空集 $W=\{x_i+S_i\mid i\in K\}$ 的交集要么是空集,要么是一个平坦集. 如果这一交集是非空的,那么对于交集中任意向量 x,

$$\bigcap_{i\in K}(x_i+S_i)=x+\bigcap_{i\in K}S_i.$$

证明 如果 $x\in\bigcap_{i\in K}(x_i+S_i)$,那么对于 $\forall i\in K, x_i+S_i=x+S_i$,所以

$$\bigcap_{i\in K}(x_i+S_i)=\bigcap_{i\in K}(x+S_i)=x+\bigcap_{i\in K}S_i.$$

定义 14.3.1 V 中平坦集所成的一非空集 $W=\{x_i+S_i\mid i\in K\}$ 的**并** Ω 是 W 中最小的一个平坦集. 我们把平坦集所成的集合 W 的并集记为 $\vee W$,或 $\bigvee_{i\in K}\{x_i+S_i\}$. 两个平坦集的并记为$(x+S)\vee(y+T)$.

定理 14.3.2 令 $W=\{x_i+S_i\mid i\in K\}$ 是 V 中平坦集所成的一非空集.

1) $\vee W$ 是包含 W 中全体平坦集的所有平坦集的交.

2) $\vee W$ 即为 $\mathrm{hull}(W)$,其中 Ω 是 W 中全体平坦集的并.

定理 14.3.3 对于 V 中任意两个平坦集,$(x+S)\vee(y+T)=x+[\langle x-y\rangle+$

$S+T]$.

证明　由于 $x, y \in (x+S) \vee (y+T)$，所以对于 V 的任意子空间 U，

$$(x+S) \vee (y+T) = x+U = y+U.$$

因此，$x-y \in U$，所以 $\langle x-y \rangle \subset U$. 而且，由 $x+S \subset x+U$ 可得 $S \subset U$.

同理，$T \subset U$. 从而，$[\langle x-y \rangle + S + T] \subset x+U$. 由于 $x+S$ 和 $y+T$ 都包含在 $x+[\langle x-y \rangle + S + T]$ 中，所以我们推出 $x+U \subset x+[\langle x-y \rangle + S + T]$.

现在，我们可以描述两个平坦集的并的维数了.

定理 14.3.4　设 $X = x+S, Y = y+T$ 是 V 的平坦集.

1) 如果 $X \cap Y \neq \varnothing$，那么

a) $X \vee Y = x+S+T$；

b) $\dim(X \vee Y) = \dim(S+T) = \dim(X) + \dim(Y) - \dim(X \cap Y)$.

2) 如果 $X \cap Y = \varnothing$，那么 $\dim(X \vee Y) = \dim(S+T)+1$.

证明　由定理 14.3.3，我们有

$$\begin{aligned}(x+S) \cap (y+T) \neq \varnothing &\Leftrightarrow \exists s \in S, t \in T, \text{使得 } x+s = y+t \\ &\Leftrightarrow x-y \in S+T \\ &\Leftrightarrow \langle x-y \rangle + S + T = S+T \\ &\Leftrightarrow X \vee Y = S+T.\end{aligned}$$

论断 1a) 和 2) 获证. 至于论断 1b)，注意到

$$\dim(S+T) = \dim(X) + \dim(Y) - \dim(X \cap Y),$$

而且，如果 $(x+S) \cap (y+T) \neq \varnothing$，那么

$$\dim(S \cap T) = \dim(x+[S \cap T]) = \dim([x+S] \cap [y+T]) = \dim(X \cap Y).$$

14.3.2　仿射无关性

现在我们讨论线性无关的对应 —— 仿射无关.

定理 14.3.5　设 $x_1, \cdots, x_n$ 是 V 中的向量，那么以下各条等价：

1) $X = \mathrm{hull}\{x_1, \cdots, x_n\}$ 有维数 $n-1$.

2) 对于 $\forall i = 1, \cdots, n, \{x_1 - x_i, \cdots, x_{i-1} - x_i, x_{i+1} - x_i, \cdots, x_n - x_i\}$ 是线性无关的.

3) 对于 $\forall i = 1, \cdots, n, x_i \notin \mathrm{hull}\{x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n\}$.

4) 如果 $\sum r_j x_j$ 和 $\sum s_j x_j$ 是仿射组合，那么对于 $\forall j, \sum_j r_j x_j = \sum_j s_j x_j \Rightarrow r_j = s_j$.

证明　从定理 14.2.1 中可以直接推出 1) 和 2) 等价.

1)⇒3)　如果 3) 不成立，那么

$$\operatorname{hull}\{x_1, \cdots, x_n\} = \operatorname{hull}\{x_1, \cdots, x_{i-1}, x_{i+1}, \cdots, x_n\}.$$

根据(14.5)式,后者的维数最多为 $n-2$. 因此 1) 不成立.

3)⇒4)　设论断 3) 成立,且 $\sum r_j x_j = \sum s_j x_j$. 规定 $t_j = r_j - s_j$,得

$$\sum_j t_j x_j = 0, \sum_j t_j = 0.$$

但如果任意 t_j 不为 0,比如 $t_1 \neq 0$,那么除以 t_1 得

$$x_1 + \sum_{j>1}\left(\frac{t_j}{t_1}\right)x_j = 0,\text{或 } x_1 = \sum_{j>1} -\left(\frac{t_j}{t_1}\right)x_j,\text{这里}\sum_{j>1} -\left(\frac{t_j}{t_1}\right)x_j = 1.$$

因此 $x_1 \in \operatorname{hull}\{x_1, \cdots, x_n\}$. 这一矛盾说明对于 $\forall j, t_j = 0$,即对于 $\forall j, r_j = s_j$.

4)⇒2)　为了更具体,由 4) 可得$\{x_2 - x_1, \cdots, x_n - x_1\}$ 是线性无关的.

事实上,如果 $\alpha_1, \cdots, \alpha_n \in F, \sum \alpha_j = \alpha$,那么

$$\sum_{j\geqslant 2}\alpha_j(x_j - x_1) = 0 \Rightarrow \sum_{j\geqslant 2}\alpha_j x_j = \alpha x_1 \Rightarrow (1-\alpha)x_1 + \sum_{j\geqslant 2}\alpha_j x_j = x_1.$$

这是两个仿射组合之间的等式,所以对应系数必相等,则对于 $\forall j = 2, \cdots, n, \alpha_j = 0$.

定义 14.3.2　向量 $x_1, \cdots, x_n$ 是**仿射无关**的,如果它们满足定理 14.3.5 的任意(因而所有)条件.

定理 14.3.6　如果 X 是一维数为 n 的平坦集,那么存在 $n+1$ 个向量 $x_1, \cdots, x_{n+1}$,使得每个向量 $x \in X$ 都有唯一的作为一仿射组合的表达式 $x = r_1 x_1 + \cdots + r_{n+1} x_{n+1}$. 系数 r_i 称为 x 关于向量 $x_1, \cdots, x_{n+1}$ 的重心坐标.

14.4　仿射变换与射影几何

14.4.1　仿射性

现在我们讨论一下保持仿射结构的映射的一些性质.

定义 14.4.1　保持仿射组合,即使得

$$\sum_i r_i = 1 \Rightarrow f\left(\sum_i r_i x_i\right) = \sum_i r_i f(x_i)$$

的函数 $f: V \to V$ 叫做**仿射变换**(或**仿射映射**,或**仿射性**).

为了使 f 为一个仿射变换,一些作者要求 f 是一个双射函数. 以下是定理 14.2.2 的类似定理.

定理 14.4.1　如果 $\operatorname{char}(F) \neq 2$,那么对于函数 $f: V \to V$,以下各条等价:

1) f 保持其任意两点的仿射组合,即 $f(rx + (1-r)y) = rf(x) + (1-r)f(y)$.

2) f 保持仿射组合,即 $\sum_i r_i = 1 \Rightarrow f(\sum_i r_i x_i) = \sum_i r_i f(x_i)$.

因此,如果 $\mathrm{char}(F) \neq 2$,那么映射 f 是一个仿射变换,当且仅当 f 把穿过 x 和 y 的直线映到穿过 $f(x)$ 和 $f(y)$ 的直线上.显然,线性变换是仿射变换.

定义 14.4.2　令 $v \in V$.对于 $\forall x \in V$,由 $T_V(x) = x + v$ 定义的仿射映射 T_V: $V \to V$ 叫做 ***v* 平移**.

不难看出,形为 $T_V \circ \tau$ 的任意映射是仿射映射,这里 $\tau \in L(V)$.反之,任意仿射映射必有这一形式.

定理 14.4.2　函数 f: $V \to V$ 是一个仿射变换 $\Leftrightarrow f = T_V \circ \tau$,这里 $v \in V, \tau \in L(V)$.

证明　设 f 是一个仿射映射,那么

$$f(rx + sy) = f[rx + sy + (1 - r - s)0] = rf(x) + sf(y) + (1 - r - s)f(0),$$

重新整理得

$$f(rx + sy) - f(0) = r[f(x) - f(0)] + s[f(y) - f(0)],$$

等价于

$$(T_{-f(0)} \circ f)(rx + sy) = r(T_{-f(0)} \circ f)(x) + s(T_{-f(0)} \circ f)(y).$$

所以 $\tau = T_{-f(0)} \circ f$ 是线性的.从而,$f = T_{-f(0)} \circ \tau$.

推论 14.4.3　1) 两个仿射变换的合成是一个仿射变换.

2) 仿射变换 $f = T_V \circ \tau$ 是双射的,当且仅当 τ 是双射的.

3) V 上全体双射仿射变换所成的集合 $\mathrm{Aff}(V)$ 在映射的合成之下是一个群,叫做 V 的**仿射群**.

以下是大家所熟知的群理论的基本性质.V 的全体平移所成集合 $\mathrm{Trans}(V)$ 是 $\mathrm{Aff}(V)$ 的一个子群.由 $\varphi(T_V \circ \tau) = \tau$,我们可以定义函数 φ: $\mathrm{Aff}(V) \to L(V)$.不难看出,φ 是从 $\mathrm{Aff}(V)$ 到 $L(V)$ 的一个唯一确定的群同态,核为 $\mathrm{Trans}(V)$.因此,$\mathrm{Trans}(V)$ 是 $\mathrm{Aff}(V)$ 的一个正规子群,且

$$\frac{\mathrm{Aff}(V)}{\mathrm{Trans}(V)} \approx L(V).$$

14.4.2　射影几何

如果 $\dim(V) = 2$,那么 V 中任意两个互异点的并是一条直线.另一方面,任意两条直线的交并不是一个点.因而,在 V 中点和直线的概念之间存在着某种非对称性.通过构造所谓的射影平面,可以去掉这种非对称性.我们这儿的平面只是简要地描述一下任意维数的射影几何的一种可能构造(图 14-1)

H 是 3 维向量空间 V 的一超平面,$0 \notin H$.现在,H 中 V 的所有平坦集所成的集合

$\mathcal{A}(H)$ 是维数为 2 的一个仿射几何.

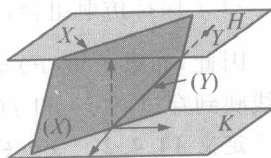


图 14－1

根据仿射几何的定义，为了定义$\mathcal{A}(H)$，H 必为一个向量空间. 然而，为了包含 V 中所有平坦集所成的集合，对于 H 中每一个平坦集 X，我们使 V 的由 X 生成的子空间 $< X >$ 相伴. 这就定义了一个函数

$$P: \mathcal{A}(H) \to \mathcal{S}(H), \quad P(X) = < X >.$$

这里$\mathcal{S}(V)$ 是V 的所有子空间所成的集合. P 不映到$\mathcal{S}(V)$ 上，因为 $\mathrm{Im}(P)$ 不包含子空间 K 的任意子空间，其中 K 中含有原点，且平行于 H. 图 14－1 展示了一个一维平坦集 X 和它的象 $P(X) = < X >$，以及一个零维平坦集 Y 和它的象 $< Y >$.

注意，对于 H 中任意平坦集 X，我们都有 $\dim(P(X)) = \dim(X) + 1$.

定义 14.4.3　令V 是一个向量空间. V 的全体子空间所成集合$\mathcal{P}(V)$ 叫做V 的**射影几何**. 如果 S 是 V 的一个子空间，那么它的**射影维数**等于 $\dim(S) - 1$，记为 $\mathrm{pdim}(S)$. $\mathcal{P}(V)$ 的**射影维数**定义为 $\mathrm{pdim}(V) = \dim(V) - 1$. 射影维数为 0，1 或 2 的子空间分别叫做**射影点**、**射影直线**或**射影平面**.

14.4.3　从仿射几何到射影几何

因此，在图 14－1 中，射影点是一条穿过原点的直线，而且，如果它不包含在刚才所说的平面 K 中，那么它在一个(仿射) 点处与 H 相交.

同理，一条射影直线是一个穿过原点的平面，假如它不是 K，那么它就与 H 相交于一条直线(这在较高的维数中也成立).

简言之，P(点) = 射影点，P(直线) = 射影直线，P(平面) = 射影平面等等.

给定一个任意维数的向量空间V，以及V 中不包含原点的任意超平面H，像在图 14－1 中 $\dim(V) = 3$ 时一样，我们可以定义函数 $P: \mathcal{A}(H) \to \mathcal{P}(V)$. 显然，从图中还可以看出，射影维数为 n 的射影几何是(仿射) 维数为 n 的一个仿射几何的“扩充”，以这种方式形成的扩充使得所有“对象” 都相交. 更具体一点，映射 $P: \mathcal{A}(H) \to \mathcal{P}(V)$ 满足以下定理所描述的性质.

定理 14.4.4　从仿射几何$\mathcal{A}(H)$ 到射影几何$\mathcal{P}(V)$ 的映射 $p: \mathcal{A}(H) \to \mathcal{P}(V)$ 满足

1) p 是单射的，其逆由 $p^{-1}(U) = U \cap H$ 得出.

2) $\mathrm{Im}(p)$ 是 V 的所有子空间的集合，其中这一集合不包含在平行于 H 的子空间K 中.

3) $X \subset Y$，当且仅当 $p(X) \subset p(Y)$.

4) 如果 X_i 是 H 中有非空交集的平坦集，那么 $p(\bigcap_{i\in K} X_i) = \bigcap_{i\in K} p(X_i)$.

5) 对于 H 中平坦集所成的任意集合，$p(\bigvee_{i\in K} X_i) = \bigoplus_{i\in K} p(X_i)$.

6) p 保持维数，在这个意义上，$\mathrm{pdim}(p(X)) = \dim(X)$.

7) $X \parallel Y \Leftrightarrow p(X) \cap K$ 包含在 $p(Y) \cap K$ 中，或 $p(Y) \cap K$ 包含在 $p(X) \cap K$ 中.

证明　1) 设 $x + S$ 是 H 中一平坦集，那么 $x \in H$，所以 $H = x + K$ 蕴涵着 $S \subset K$. 同时，$p(x+S) = \langle x \rangle + S$，且对于 $\forall s \in S, k \in K$ 和 $r \in F$，

$$z \in p(x+S) \cap H = (\langle x \rangle + S) \cap x + K \Rightarrow z = rx + s = x + k,$$

则 $(1-r)x \in K$，从而 $x \in K$ 或 $r = 1$. 但 $x \in H$ 可得 $x \notin K$，所以 $r = 1$，则 $z = x + s \in x + S$，即 $p(x+S) \cap H \subset x + S$. 由于反包含是显然的，所以 $p(x+S) \cap H = x + S$.

2) 设 U 是 V 的不包含在 K 中的子空间. 我们想要说明 U 在$\mathcal{P}$的象中. 首先，由于 $U \not\subset K, \dim(K) = \dim(V) - 1$，所以 $U + K = V$，从而

$$\dim(U \cap K) = \dim(U) + \dim(K) - \dim(U + K) = \dim(U) - 1.$$

现在，设 $0 \neq x \in U - K$，那么

$$x \notin K \Rightarrow \langle x \rangle + K = V \Rightarrow \text{对于 } \forall 0 \neq r \in F, k \in K, rx + k \in H \Rightarrow rx \in H.$$

因此对于 $\forall 0 \neq r \in F, rx \in U \cap H$，从而平坦集 $rx + (U \cap H)$ 在 H 中，且

$$\dim(rx + (U \cap H)) = \dim(U \cap K) = \dim(U) - 1,$$

则

$$p(rx + (U \cap H)) = \langle rx \rangle + (U \cap K)$$

在 U 中，且与 U 有相同维数，即

$$p(rx + (U \cap H)) = \langle rx \rangle + (U \cap K) = U.$$

14.5　形式幂级数

14.5.1　形式幂级数

定义 14.5.1　令$\mathcal{F}$表示变量 t 中有复系数的形式幂级数的代数. 因此$\mathcal{F}$是形为

$$f(t) = \sum_{k=0}^{\infty} a_k t^k$$

的全体形式和所成的集合，这里 $a_k \in \mathbf{C}$. 加法和乘法是纯形式

$$\sum_{k=0}^{\infty} a_k t^k + \sum_{k=0}^{\infty} b_k t^k = \sum_{k=0}^{\infty} (a_k + b_k) t^k, \left(\sum_{k=0}^{\infty} a_k t^k\right)\left(\sum_{k=0}^{\infty} b_k t^k\right) = \sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} a_j b_{k-j}\right) t^k.$$

f 的阶 $o(f)$ 是 t 的有非零系数的最小指数. 零级数的阶是 $+\infty$，级数 f 有一个乘法逆，记为 f^{-1}，当且仅当 $o(f) = 0$，而且

$$o(fg)=o(f)+o(g),o(f+g)\geqslant \min\{o(f),o(g)\}.$$

如果 f_k 是当$\mathcal{F}$中 $k\to 0$ 时,$o(f_k)\to\infty$ 的一个序列,那么对于任意级数 $g(t)=\sum_{k=0}^{\infty}b_kt^k$,我们都可以构成级数 $h(t)=\sum_{k=0}^{\infty}b_kf_k(t)$. 这个和是唯一确定的,因为 t 的每个幂的系数都是一个有限和.

特别地,如果 $o(f)\geqslant 1$,那么 $o(f_k)\to\infty$,所以合成

$$(g\circ f)(t)=g(f(t))=\sum_{k=0}^{\infty}b_kf^k(t)$$

是唯一确定的. 容易看出 $o(g\circ f)=o(g)o(f)$.

定义 14.5.2 如果 $o(f)=1$,那么 f 有一个合成逆,记为 $\overline{f}$. 满足

$$(f\circ\overline{f})(t)=(\overline{f}\circ f)(t)=t. o(f)=1$$

的级数 f 叫做 **δ 级数**.

δ 级数 f 的幂 f^k 所成序列构成$\mathcal{F}$的一个伪基,在这个意义上,对于 $\forall g\in\mathcal{F}$,都存在唯一的常数序列 a_k,使得 $g(t)=\sum_{k=0}^{\infty}a_kf^k(t)$. 级数 $f(t)=\sum_{k=0}^{\infty}a_kt^k$ 的形式导数由 $\partial_tf(t)=f'(t)=\sum_{k=1}^{\infty}ka_kt^{k-1}$ 给出. 算子 ∂_t 是一个导子,即 $\partial_t(fg)=\partial_t(f)g+f\partial_t(g)$.

14.5.2 幂级数和线性泛函的关系

令$\mathcal{P}=\mathbf{C}[x]$ 表示复数域上单变量 x 的多项式代数. $\mathcal{F}$中任意形式幂级数都能扮演三种角色 —— 作为一个形式幂级数、作为$\mathcal{P}$上一个线性泛函,以及作为$\mathcal{P}$上一个线性算子. 我们首先讨论形式幂级数和线性泛函之间的关系.

设$\mathcal{P}^*$ 表示$\mathcal{P}$上全体线性泛函所成的向量空间. $\mathcal{P}^*$ 是$\mathcal{P}$的代数对偶空间. 由 $<L\mid p(x)>$ 将方便地在 $p(x)\in\mathcal{P}$上表示出 $L\in\mathcal{P}^*$ 的作用,于是$\mathcal{P}^*$ 上的向量空间运算取形式

$$<L+M\mid p(x)>=<L\mid p(x)>+<M\mid p(x)>;$$
$$<rL\mid p(x)>=r<L\mid p(x)>,r\in\mathbf{C}.$$

由于$\mathcal{P}$上任意线性泛函都是由$\mathcal{P}$的一个基的值唯一确定的,所以对于 $n\geqslant 0$,$L\in\mathcal{P}^*$ 由其值 $<L\mid x^n>$ 唯一确定.

$\mathcal{F}$中任意形式级数都可以记为形式 $f(t)=\sum_{k=0}^{\infty}\frac{a_k}{k!}t^k$. 对于 $n\geqslant 0$,规定 $<f(t)\mid x^n>=a_n$,用这一形式来定义线性泛函 $f(t)$,即线性泛函 $f(t)$ 由条件

$$f(t)=\sum_{k=0}^{\infty}\frac{< f(t)\mid x_k >}{k!}t^k$$

定义. $< t^k \mid x^n > = n!\delta_{n,k}$，这里 $\delta_{n,k}$ 是克罗内克δ 函数，则 $< t^k \mid p(x) > = p^{(k)}(0)$，所以 t^k 是函数的"在 0 上的第 k 个导数". 同样，t^0 是在 0 处的求值.

巧合的是，任意线性泛函 $L \in \mathcal{P}^*$ 都有形式 $f(t)$. 为了说明这一点，如果

$$f_L(t)=\sum_{k=0}^{\infty}\frac{< L \mid x_k >}{k!}t^k,$$

那么对于 $\forall n \geqslant 0$，$< f_L(t) \mid x^n > = < L \mid x^n >$，所以作为一个线性泛函，$L = f_L(t)$.

因此，由 $\varphi(L) = f_L(t)$，我们可以定义映射 $\varphi\colon \mathcal{P}^* \to \mathcal{F}$.

定理 14.5.1　由 $\varphi(L) = f_L(t)$ 定义的映射 $\varphi\colon \mathcal{P}^* \to \mathcal{F}$是从$\mathcal{P}^*$ 映到$\mathcal{F}$上的一个向量空间同构.

证明　为了说明 φ 是单射的，注意到

$$f_L(t) = f_M(t) \Rightarrow \text{对于 } \forall n \geqslant 0,\ < L \mid x^n > = < M \mid x^n > \Rightarrow L = M.$$

而且映射 φ 是满射的，因为对于任意 $f \in \mathcal{F}$，线性泛函 $L = f(t)$ 都满足 $\varphi(L) = f_L(t) = f(t)$.

最后，

$$\varphi(rL+sM)=\sum_{k=0}^{\infty}\frac{< rL+sM \mid x_k >}{k!}t^k = r\sum_{k=0}^{\infty}\frac{< L \mid x_k >}{k!}t^k + s\sum_{k=0}^{\infty}\frac{< M \mid x_k >}{k!}t^k$$
$$= r\varphi(L) + s\varphi(M).$$

现在，我们就可以用同构 $\varphi\colon \mathcal{P}^* \to \mathcal{F}$来把向量空间$\mathcal{P}^*$ 和向量空间$\mathcal{F}$视为一体了. 从而，我们只要简单地把$\mathcal{P}^*$ 上的线性泛函看作是形式幂级数. 这种方法的优点是$\mathcal{F}$不只是一个向量空间$\mathcal{F}$，它还是一个代数. 因此，我们就定义了线性泛函的一个乘法，即形式幂级数的积. 代数$\mathcal{F}$既可以看成是形式幂级数的代数，又可以看成是$\mathcal{P}$上线性泛函的代数.

例 14.5.1　对于 $a \in \mathbf{C}$，**求值泛函** $\varepsilon_a \in \mathcal{P}^*$ 由 $< \varepsilon_a \mid p(x) > = p(a)$ 定义.

特别地，$< \varepsilon_a \mid x^n > = a^n$，所以这个泛函的形式幂级数表示为

$$f_{\varepsilon_a}(t)=\sum_{k=0}^{\infty}\frac{< \varepsilon_a \mid x^k >}{k!}t^k=\sum_{k=0}^{\infty}\frac{a^k}{k!}t^k=\mathrm{e}^{at}.$$

它是一个指数级数. 如果 e^{bt} 是在b 处求值，那么 $\mathrm{e}^{at}\mathrm{e}^{bt} = \mathrm{e}^{(a+b)t}$. 所以在 a 处求值和在b 处求值的积是在 $a+b$ 处求值.

当把一个 δ 级数 $f \in \mathcal{F}$看成是一个线性泛函时，我们称 f 为一个 **δ 泛函**. 同样，可逆级数 $f \in \mathcal{F}$称为是一个**可逆泛函**.

以下是我们所得出的一些简单结论：

定理 14.5.2 1）对于 $\forall f \in F, f(t) = \sum_{k=0}^{\infty} \frac{< f(t) \mid x^k >}{k!} t^k$.

2）对于 $\forall p \in \mathcal{F}, p(x) = \sum_{k \geqslant 0} \frac{< t^k \mid p(x) >}{k!} x^k$.

3）对于 $\forall f, g \in \mathcal{F}, < f(t)g(t) \mid x^n > = \sum_{k=0}^{n} \binom{n}{k} < f(t) \mid x^k >< g(t) \mid x^{n-k} > t^k$.

4）$o(f(t)) > \deg p(x) \Rightarrow < f(t) \mid p(x) > = 0$.

5）如果对于 $\forall k \geqslant 0, o(f_k) = k$，那么

$$< \sum_{k=0}^{\infty} a_k f_k(t) \mid p(x) > = \sum_{k \geqslant 0} a_k < f_k(t) \mid p(x) >.$$

这里等式右边的和是一个有限和.

6）如果对于 $\forall k \geqslant 0, o(f_k) = k$，那么对于 $\forall k \geqslant 0$，

$$< f_k(t) \mid p(x) > = < f_k(t) \mid q(x) > \Rightarrow p(x) = q(x).$$

7）如果对于 $\forall k \geqslant 0, \deg p_k(x) = k$，那么对于 $\forall k \geqslant 0$，

$$< f(t) \mid p_k(x) > = < g(t) \mid p_k(x) > \Rightarrow f(t) = g(t).$$

证明 我们只证论断 3）. 令 $f(t) = \sum_{k=0}^{\infty} \frac{a_k}{k!} t^k, g(t) = \sum_{j=0}^{\infty} \frac{b_j}{j!} t^j$，那么

$$f(t)g(t) = \sum_{m=0}^{\infty} \left(\frac{1}{m!} \sum_{k=0}^{m} \binom{m}{k} a_k b_{m-k} \right) t^m,$$

等式两边（作为线性泛函）应用 x^n，得

$$< f(t)g(t) \mid x^n > = \sum_{k=0}^{n} \binom{n}{k} a_k b_{n-k}.$$

由于论断 1）有 $a_k = < f(t) \mid x^k >, b_{n-k} = < g(t) \mid x^{n-k} >$，由此获证.

定理 14.5.3 对于 $\forall f(t) \in \mathcal{F}, p(x) \in \mathcal{P}, < f(t) \mid xp(x) > = < \partial_t f(t) \mid p(x) >$.

证明 由线性性，只要证 $p(x) = x^n$ 时的情形. 但是，如果 $f(t) = \sum_{k=0}^{\infty} \frac{a_k}{k!} t^k$，那么

$$< \partial_t f(t) \mid x^n > = < \sum_{k=1}^{\infty} \frac{a_k}{(k-1)!} t^{k-1} \mid x^n > = \sum_{k=1}^{\infty} \frac{a_k}{(k-1)!} \delta_{k-1,n} = a_{n+1} = < f(t) \mid x^{n+1} >.$$

14.5.3 线性泛函的幂级数表示

我们考虑以下几个重要的线性泛函及它们的幂级数表示的例子.

例 14.5.2　1）**求值泛函** e^{at} 满足 $< e^{at} \mid p(x) >= p(a)$.

2）**前向差分泛函**是 δ 泛函 $e^{at}-1$，满足 $< e^{at}-1 \mid p(x) >= p(a)-p(0)$.

3）**阿贝尔泛函**是 δ 泛函 te^{at}，满足 $< te^{at} \mid p(x) >= p'(a)$.

4）**可逆泛函** $(1-t)^{-1}$ 满足 $<(1-t)^{-1} \mid p(x) >= \int_0^{\infty} p(u)e^{-u}du$.

通过令 $p(x)=x^n$，同时扩充表达式 $(1-t)^{-1}$，可以得出上述等式.

5）为了确定线性泛函 f 满足 $< f(t) \mid p(x) >= \int_0^a p(u)du$，注意到

$$f(t)=\sum_{k=0}^{\infty}\frac{< f(t) \mid x^k >}{k!}t^k=\sum_{k=0}^{\infty}\frac{a^{k+1}}{(k+1)!}t^k=\frac{e^{at}-1}{t}.$$

这一泛函的逆 $\dfrac{t}{e^{at}-1}$ 与伯努利多项式有关，后者无论是在数学上，还是在应用上都起着重要的作用.

事实上，$B_n=< \dfrac{t}{e^{at}-1} \mid x^n >$ 就是**伯努利数**.

14.6　几种重要的线性算子和多项式

14.6.1　线性算子的形式幂级数

现在我们讨论形式幂级数与 $\mathcal{P}$ 上线性算子的联系. 用 t^k 表示 $\mathcal{P}$ 上第 k 个导数算子，因此 $t^k p(x)=p^{(k)}(x)$. 把它扩充为 t 中的形式幂级数

$$f(t)=\sum_{k=0}^{\infty}\frac{a^k}{k!}t^k. \tag{14.6}$$

通过 $f(t)p(x)=\sum_{k=0}^{\infty}\dfrac{a^k}{k!}[t^k p(x)]=\sum_{k\geqslant 0}\dfrac{a^k}{k!}p^{(k)}(x)$ 来定义线性算子 $f(t)$：$\mathcal{P}\to P$.

后面的和是一个有限和

$$f(t)x^n=\sum_{k=0}^{\infty}\binom{n}{k}a_k x^{n-k}. \tag{14.7}$$

根据这个定义，我们看到每个形式幂级数 $f\in\mathcal{F}$ 都扮演着三种角色，即，作为一个形式幂级数、作为一个线性泛函，以及作为一个线性算子. 不同的符号 $< f(t) \mid p(x) >$ 和 $f(t)p(x)$ 说明了我们把 f 看成是一个泛函还是一个算子.

在 $\mathcal{F}$ 中，$f=g$，当且仅当 $f=g$ 是线性泛函. 而这一表述成立，当且仅当 $f=g$ 又是线性算子.

同样，$[f(t)g(t)]p(x)=f(t)[g(t)p(x)]$. 所以我们可以记作 $f(t)g(t)p(x)$.

另外，对于 $\forall f,g\in\mathcal{F},p\in\mathcal{P},f(t)g(t)p(x)=g(t)f(t)p(x)$. 当把 δ 级数 f 看成一个算子时，我们称 f 为 **δ 算子**.

以下定理叙述了线性泛函与形为 $f(t)$ 的线性算子之间的重要关系.

定理 14.6.1 如果 $f,g\in\mathcal{F}$，那么对于任意多项式 $p(x)\in\mathcal{P}$，

$$\langle f(t)g(t)\mid p(x)\rangle=\langle f(t)\mid g(t)p(x)\rangle.$$

证明 如果 f 有形式(14.6)，那么根据(14.7)式，

$$\langle t^0\mid f(t)x^n\rangle=\langle t^0\mid\sum_{k=0}^{n}C_n^k a_k x^{n-k}\rangle=a_n=\langle f(t)\mid x^n\rangle. \tag{14.8}$$

由此得 $\langle f(t)g(t)\mid p(x)\rangle=\langle t^0\mid f(t)g(t)p(x)\rangle=\langle t^0\mid f(t)[g(t)p(x)]\rangle=\langle f(t)\mid g(t)p(x)\rangle$.

等式(14.8)表明，应用线性泛函 $f(t)$ 与应用算子 $f(t)$ 是等价的，通过在 $x=0$ 处求值可以得出这一结论.

例 14.6.1 1）算子 e^{at} 满足 $e^{at}x^n=\sum\limits_{k=0}^{\infty}\dfrac{a^k}{k!}t^kx^n=\sum\limits_{k=0}^{n}C_n^ka_kx^{n-k}=(x+a)^n$. 所以，对于 $\forall p\in\mathcal{P},e^{at}p(x)=p(x+a)$，从而 e^{at} 是一个**平移算子**.

2）**前向差分算子**是 δ 算子 $e^{at}-1$，其中 $(e^{at}-1)p(x)=p(x+a)-p(a)$.

3）**阿贝尔算子**是 δ 算子 te^{at}，其中 $te^{at}p(x)=p'(x+a)$.

4）**可逆算子** $(1-t)^{-1}$ 满足 $(1-t)^{-1}p(x)=\int_0^{\infty}p(x+u)e^{-u}du$.

5）易见算子 $(e^{at}-1)/t$ 满足 $\dfrac{e^{at}-1}{t}p(x)=\int_x^{x+a}p(u)du$.

对于 $f\in\mathcal{F}$，$\mathcal{P}$上所有线性泛函都有 $f(t)$ 形式. 但并不是$\mathcal{P}$上所有线性泛函都有这种形式. 事实上，注意到 $\deg[f(t)p(x)]\leqslant\deg p(x)$. 但由 $\varphi(p(x))=xp(x)$ 定义的线性算子并没有这一性质.

定理 14.6.2 对于线性算子 $\tau:\mathcal{P}\to\mathcal{P}$，以下各条等价：

1）τ 有形式 $f(t)$，即存在一作为线性算子的 $f\in\mathcal{F}$，使得 $\tau=f(t)$.

2）τ 可与导数算子交换，即 $\tau t=t\tau$.

3）τ 可与任意 δ 算子 $g(t)$ 交换，即 $\tau g(t)=g(t)\tau$.

4）τ 可与任意平移算子交换，即 $\tau e^{at}=e^{at}\tau$.

14.6.2 谢弗尔序列

如果提到$\mathcal{P}$中的序列 $s_n(x)$，那么就有对于 $\forall n\geqslant 0,\deg s_n(x)=n$.

定理 14.6.3 令 f 是一个 δ 级数，g 是一可逆级数，考虑$\mathcal{F}$中等比数列 g,gf,gf^2，

gf^3,…,那么在$\mathcal{P}$中存在唯一的序列 $s_n(x)$,对于 $\forall n,k \geqslant 0$,满足正交性条件

$$< g(t)\ f^k(t) \mid s_n(x) >= n!\delta_{n,k}.$$

定义 14.6.1　此式中的序列 $s_n(x)$ 称为有序对$(g(t),f(t))$ 的**谢弗尔序列**.

以下是谢弗尔序列的两种特殊类型:

定义 14.6.2　形为 $(1,f(t))$ 的有序对的谢弗尔序列称为 $f(t)$ 的**相伴序列**. 形为 $(g(t),t)$ 的有序对的谢弗尔序列称为 $g(t)$ 的 **Abel 序列**.

定理 14.6.4(展开定理)　设 $s_n(x)$ 是有序对$(g(t),f(t))$ 的谢弗尔序列.

1) 对于 $\forall h \in \mathcal{F}, h(t) = \sum\limits_{k=0}^{\infty} \dfrac{< h(t) \mid s_k(x) >}{k!} g(t) f^k(t)$.

2) 对于 $\forall p \in \mathcal{P}, p(x) = \sum\limits_{k\geqslant 0} \dfrac{< g(t)f^k(t) \mid p(x) >}{k!} s_k(x)$.

证明　论断 1) 可以从定理 14.5.2 中的 5) 和 7) 得到,因为

$$< \sum_{k=0}^{\infty} \frac{< h(t) \mid s_k(x) >}{k!} g(t) f^k(t) \mid s_n(x) > = \sum_{k=0}^{\infty} \frac{< h(t) \mid s_k(x) >}{k!} n!\delta_{n,k}$$

$$= < h(t) \mid s_n(x) >.$$

类似地,论断 2) 可以从定理 14.5.2 中的 6) 得出.

14.6.3　谢弗尔序列的特征

现在,我们可以描述谢弗尔序列的特征了. 首先是母函数. 母函数的思想很简单. 如果 $r_n(x)$ 是一多项式序列,那么可以定义一个形为 $g(t,x) = \sum\limits_{k=0}^{\infty} \dfrac{r_k(x)}{k!} t^k$ 的形式幂级数. 这称为序列 $r_n(x)$ 的**(指数)母函数**. 由于这个级数是形式级数,所以知道了 $g(t)$ 就等于知道了多项式 $r_n(x)$. 而且,了解一个多项式序列的母函数,我们会对这个序列本身有更深的理解,而这用其他途径可能是达不到的. 由于这一原因,对母函数的研究相当重要.

定理 14.6.5　1) 令 $P_n(x)$ 是 $f(t)$ 的相伴序列,$\bar{f}(t)$ 是 $f(t)$ 的合成逆,则 $P_n(x)$ 的母函数是 $\mathrm{e}^{y\bar{f}(t)} = \sum\limits_{k=0}^{\infty} \dfrac{P_k(y)}{k!} t^k$.

2) 令 $s_n(x)$ 是有序对$(g(t),f(t))$ 的谢弗尔序列,$s_n(x)$ 的母函数是

$$\frac{1}{g(\bar{f}(t))} \mathrm{e}^{y\bar{f}(t)} = \sum_{k=0}^{\infty} \frac{s_k(y)}{k!} t^k.$$

定理 14.6.6　1) $P_n(x)$ 是 $f(t)$ 的相伴序列 $\Leftrightarrow P_n(x) = \sum\limits_{k=0}^{n} \dfrac{1}{k!} < \bar{f}(t)^k \mid x^n > x^k$.

2) $s_n(x)$ 是 $(g(t), f(t))$ 的谢弗尔序列 $\Leftrightarrow s_n(x) = \sum_{k=0}^{n} \frac{1}{k!} < g(\bar{f}(t))^{-1} \bar{f}(t)^k \mid x^n > x^k$.

定理 14.6.7（算子特征） 1) 序列 $P_n(x)$ 是 $f(t)$ 的相伴序列，当且仅当

a) $P_n(x) = \delta_{n,0}$； b) 对于 $n \geqslant 0, f(t)P_n(x) = nP_{n-1}(x)$.

2) 对于 $\forall g(t)$，序列 $s_n(x)$ 是有序对 $(g(t), f(t))$ 的谢弗尔序列，当且仅当对于 $\forall n \geqslant 0, f(t)s_n(x) = ns_{n-1}(x)$.

定理 14.6.8 1)（**二项式恒等式**） $P_n(x)$ 是 δ 级数 $f(t)$ 的相伴序列 $\Leftrightarrow P_n(x)$ 有二项式型，即当且仅当对于 $\forall y \in \mathbf{C}, P_n(x)$ 满足恒等式 $P_n(x+y) = \sum_{k=0}^{n} C_n^k P_k(y) P_{n-k}(x)$.

2)（**谢弗尔恒等式**） 对 $\forall g(t), s_n(x)$ 是 $(g(t), f(t))$ 的谢弗尔序列 $\Leftrightarrow$ 对 $\forall y \in \mathbf{C}, s_n(x+y) = \sum_{k=0}^{n} C_n^k P_k(y) s_{n-k}(x)$，这里 $P_n(x)$ 是 $f(t)$ 的相伴序列.

14.6.4 几个重要的多项式

现在我们可以举出谢弗尔序列的一些例子了. 虽然验证给定序列是有序对 $(g(t), f(t))$ 的谢弗尔序列常常是一件相对简单的事情，但找出一给定有序对的谢弗尔序列却比较困难. 我们为找出谢弗尔序列提供了两个公式，一个是直接的，但通常需要计算级数 $(f(t)/t)^{-n}$，而这是很困难的. 另一个是递推关系，用来表示谢弗尔序列中每一个根据前一项得出的 $s_n(x)$. 这里我们只讨论相伴序列的递推公式.

例 14.6.2 序列 $P_n(x) = x^n$ 是 δ 级数 $f(t) = t$ 的相伴序列. 这一序列的母函数是 $e^{yt} = \sum_{k=0}^{\infty} \frac{y^k}{k!} t^k$. 二项式恒等式恰是 $(x+y)^n = \sum_{k=0}^{n} C_n^k x^k y^{n-k}$.

例 14.6.3 下阶乘多项式 $(x)_n = x(x-1)\cdots(x-n+1)$ **构成前向差分泛函** $f(t) = e^t - 1$ 的相伴序列.

我们用定理 14.6.7 来计算一下. 由于 $(0)_0$ 定义为 1，所以 $(0)_n = \delta_{n,0}$. 同样，

$$
\begin{aligned}
(e^t - 1)(x)_n &= (x+1)_n - (x)_n = (x+1)x(x-1)\cdots(x-n+2) - x(x-1)\cdots(x-n+1) \\
&= x(x-1)\cdots(x-n+2)[(x+1)-(x-n+1)] \\
&= nx(x-1)\cdots(x-n+2) \\
&= n(x)_{n-1}.
\end{aligned}
$$

下阶乘多项式的母函数是 $e^{y\log(1+t)} = \sum_{k=0}^{\infty} \frac{(y)_k}{k!} t^k$. 更常见的形式是 $(1+t)^y = \sum_{k=0}^{\infty} C_n^k t^k$. 这是一个形式恒等式，所以就不用在 t 上作任何限制. 在这一情况下的二项式

恒等式是$(x+y)_n=\sum_{k=0}^{n}C_n^k(x)^k(y)^{n-k}$，也可记为$C_n^{x+y}=\sum_{k=0}^{n}C_k^xC_{n-k}^y$. 这称为**范德蒙卷积公式**.

例 14.6.4　**阿贝尔多项式**$A_n(x;a)=x(x-an)^{n-1}$构成 Abel 泛函$f(t)=te^{at}$的相伴序列. Abel 多项式的母函数是$e^{y\bar{f}(t)}=\sum_{k=0}^{\infty}\frac{y(y-ak)^{k-1}}{k!}t^k$. 取它关于$y$的形式导数，得

$$\bar{f}(t)e^{y\bar{f}(t)}=\sum_{k=0}^{\infty}\frac{k(y-a)(y-ak)^{k-1}}{k!}t^k.$$

对于$y=0$，上式给出了级数$f(t)=te^{at}$的合成逆的以下一个公式：

$$\bar{f}(t)=\sum_{k=1}^{\infty}\frac{(-a)^kk^{k-1}}{(k-1)!}t^k.$$

例 14.6.5　著名的**埃尔米特多项式**$H_n(x)$构成可逆泛函$g(t)=e^{t^2/2}$的阿贝尔序列，则$s_n(x)$是$g(t)$的阿贝尔序列$\Leftrightarrow s_n(x)=g(t)^{-1}x^n$. 由此可得

$$H_n(x)=e^{-t^2/2}x^n=\sum_{k\geqslant 0}(-\frac{1}{2})^k\frac{(n)_{2k}}{k!}x^{n-k}.$$

埃尔米特多项式的母函数是$e^{yt-t^2/2}=\sum_{k=0}^{\infty}\frac{H_k(y)}{k!}t^k$. 谢弗尔恒等式是$H_n(x+y)=\sum_{k=0}^{n}C_n^kH_k(x)y^{n-k}$.

例 14.6.6　阶为α的著名**拉盖尔多项式**$L_n^{(\alpha)}(x)$构成$g(t)=(1-t)^{-\alpha-1}$，$f(t)=\frac{t}{t-1}$的谢弗尔序列. 而$L_n^{(\alpha)}(x)=\sum_{k=0}^{n}\frac{n!}{k!}C_{n-k}^{\alpha+n}(-x)^k$. 拉盖尔多项式的母函数是

$$\frac{1}{(1-t)^{\alpha+1}}e^{yt/(t-1)}=\sum_{k=0}^{\infty}\frac{L_k^{(\alpha)}(x)}{k!}t^k.$$

与埃尔米特多项式一样，拉盖尔多项式的一些定义与根据一个乘法常数得出的定义不同. 在 Roman S[1984] 中，讨论了将近 30 种不同的多项式序列，它们都是谢弗尔序列.

参 考 文 献

[1] Wan Zhe-Xian. Geometry of Classical Groups over Finite Fields: 2nd ed. Beijing: Science Press, 2002

[2] Derek J. S Robinson. A course in the Theory of Groups. New York: Springer-Verlag. 1980

[3] Roman S. Advanced Linear Algebra. New York: Springer-Verlag, 1992

[4] Jacobson N. Basic Algebra Ⅰ and Ⅱ. 2nd ed. New York: W. H. Freeman and Company, 1985 and 1989

[5] Crandall R and Pomerance C. Prime Numbers: A Computational Perspective. New York: Springer-Verlag, 2001

[6] Roman S. Coding and Information Theory. New York: Springer-Verlag, 1992

[7] Blyth T S. Medule Theory—An Approach to Linear Algebra. Oxford University. Press, 1990

[8] McEliece R J. Fields for Computer Scientisrts and Engineers. Kluwer Academic Publishers, 1987

[9] Kreyszig E. Introductory Functional Analysis with Applications. John Wiley and Sons, 1978

[10] Grueb W. Multilinear Algebra. New York: Springer-Verlag, 1978

[11] Grueb W. Linear Algebra. 4th ed. New York: Springer-Verlag, 1975

[12] Snapper E and Troyer R. Metric Affine Geometry. Dover: Springer-Verlag, 1971

[13] 刘绍学. 抽象代数. 北京：高等教育出版社，2001

[14] 张禾瑞，郝炳新. 高等代数. 第四版. 北京：高等教育出版社，1998

[15] 聂灵沼，丁石孙. 代数学引论. 北京：高等教育出版社，1988

[16] 谢邦杰. 抽象代数学. 上海：上海科学技术出版社，1982

[General Information]
书名=近世代数观点下的高等代数
作者=陈辉著
页数=320
SS号=12302130
DX号=
出版日期=2009.08
出版社=浙江大学出版社